

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 55.4461 Seconds  
(without alignments)  
3839.562 Million cell updates/sec

Title: US-10-797-893-2  
Perfect score: 1878  
Sequence: 1 MAAGSGTQORREMAASAA.....FYLDTVSALNFAARKEVIN 370

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	100.0	665	1 KF22 HUMAN	Q14807 homo sapien
2	1878	100.0	665	2 AAP35923	Aap35923 homo sapi
3	1604.5	85.4	660	2 Q991C7	Q991C7 mus musculu
4	1232	65.6	639	2 Q6P3R1	Q6P3R1 xenopus tro
5	1232	65.6	639	2 AAH63896	AAH63896 xenopus t
6	1202.5	64.0	651	2 Q91869	Q91869 xenopus lae
7	1202.5	64.0	651	2 AAH70549	AAH70549 xenopus l
8	1201.5	64.0	651	2 Q918K0	Q918K0 xenopus lae
9	1197.5	63.8	663	2 Q7ZYL5	Q7ZYL5 xenopus lae
10	1192.5	63.5	650	2 Q919A8	Q919A8 xenopus lae
11	1191.5	63.4	631	2 Q6GPG0	Q6GPG0 xenopus lae
12	951	50.6	198	2 Q60845	Q60845 homo sapien
13	704	37.5	148	2 Q35232	Q35232 mus musculu
14	623.5	33.2	584	2 Q6L512	Q6L512 oryza sativ
15	623.5	33.2	584	2 AAT39162	AAT39162 oryza sat
16	594.5	31.7	628	2 Q8L7B8	Q8L7B8 arabidopsis
17	594.5	31.7	644	2 Q9LZ88	Q9LZ88 arabidopsis
18	584.5	31.1	548	2 Q8SQ09	Q8SQ09 encephalito
19	578	30.8	912	2 Q7RX60	Q7RX60 neopospora
20	568.5	30.3	821	2 Q7QDS6	Q7QDS6 anopheles g
21	563	30.0	838	2 Q862B6	Q862B6 botrytis ci
22	559.5	29.8	784	2 Q961H5	Q961H5 drosophila
23	559	29.8	548	2 Q8N1X8	Q8N1X8 homo sapien
24	555.5	29.6	784	1 K6L8 DROME	P46867 drosophila
25	552.5	29.4	898	2 Q86V55	Q86V55 homo sapien
26	552.5	29.4	898	2 Q8N177	Q8N177 homo sapien
27	552.5	29.4	898	2 Q9H0F3	Q9H0F3 homo sapien
28	552	29.4	1296	2 Q7RM16	Q7RM16 plasmodium
29	551.5	29.4	562	2 Q8BY99	Q8BY99 mus musculu
30	551.5	29.4	571	2 Q8B245	Q8B245 mus musculu
31	551.5	29.4	880	2 Q8BL11	Q8BL11 mus musculu

32	551.5	29.4	886	2 Q91WD7	Q91WD7 mus musculu
33	551.5	29.4	895	2 Q7ZUW9	Q7ZUW9 brachydanio
34	550	29.3	703	2 Q9FZ77	Q9FZ77 arabidopsis
35	549	29.2	750	2 Q7KSK2	Q7KSK2 drosophila
36	549	29.2	750	2 AAS65150	AAS65150 drosophil
37	547	29.1	1254	2 Q94463	Q94463 dictyosteli
38	546	29.1	997	2 Q7ZXX2	Q7ZXX2 xenopus lae
39	545.5	29.0	389	2 Q7SYZ3	Q7SYZ3 xenopus lae
40	542.5	28.9	553	2 Q7P8H1	Q7P8H1 anopheles g
41	541	28.8	650	2 Q8DDE9	Q8DDE9 xenopus lae
42	536.5	28.6	1058	2 Q9LZU5	Q9LZU5 arabidopsis
43	536	28.5	784	1 KLP6 SCHPO	O59751 schizosacch
44	535.5	28.5	677	2 Q9VRK9	Q9VRK9 drosophila
45	533.5	28.4	642	2 Q7QY55	Q7QY55 giardia lam

## ALIGNMENTS

### RESULT 1

ID	KF22_HUMAN	STANDARD;	PRT;	665 AA.
AC	Q14807; Q94814; Q9BT46;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)			
DE	(Kinesin-like protein 4).			
GN	Name=KIF22; Synonyms=KNSL4, KID;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96174806; PubMed=8599929;			
RA	Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,			
RA	Tsukita S., Inoue J., Yamamoto T.,			
RT	"Kid, a novel kinesin-like DNA binding protein, is localized to			
RT	chromosomes and the mitotic spindle."			
RL	EMBO J. 15:457-467(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Lymphocytoblastoma;			
RX	MEDLINE=99009323; PubMed=9790757;			
RA	Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,			
RA	Geltinger C., Saito-Onara F., Ikeuchi T., Matsumura M., Itakura K.,			
RA	Kanazawa I., Sun K., Yokoyama K.K.;			
RT	"Human genes for KNSL4 and MAZ are located close to one another on			
RT	chromosome 16p11.2."			
RL	Genomics 52:374-377(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Brain, and Lung;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.J., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M.J., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodríguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			

This Page Blank (uspto)





This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 76.7254 Seconds  
(without alignments)  
3839.562 Million cell updates/sec

Title: US-10-797-893-4  
Perfect score: 2589  
Sequence: 1 MPAAGSTQORREMAASA.....LEAKMLAQAEKHCPTM 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_02:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2577	99.5	665	1 KF22 HUMAN	Q14807 homo sapien
2	2577	99.5	665	2 AAP35923	Aap35923 homo sapi
3	2158.5	83.4	660	2 Q991C7	Q991C7 mus musculus
4	1463	56.5	639	2 Q6P3R1	Q6P3R1 xenopus tro
5	1453	56.5	639	2 AAH63896	AAH63896 xenopus t
6	1438	55.5	651	2 Q91869	Q91869 xenopus lae
7	1438	55.5	651	2 AAH70549	AAH70549 xenopus l
8	1437	55.5	651	2 Q918K0	Q918K0 xenopus lae
9	1428	55.2	663	2 Q72VL5	Q72VL5 xenopus lae
10	1423	55.0	650	2 Q919A8	Q919A8 xenopus lae
11	1422	54.9	631	2 Q6GPG0	Q6GPG0 xenopus lae
12	946	36.5	198	2 Q60845	Q60845 xenopus lae
13	704	27.2	148	2 Q35232	Q35232 mus musculus
14	633	24.4	584	2 Q6L512	Q6L512 oryza sativ
15	633	24.4	584	2 AAT39162	AAT39162 oryza sat
16	616	23.8	784	2 Q961H5	Q961H5 drosophila
17	614	23.7	628	2 Q8L7B8	Q8L7B8 arabidopsis
18	612	23.6	664	2 Q9L288	Q9L288 arabidopsis
19	612	23.6	784	1 KL68 DROME	P46867 drosophila
20	611.5	23.6	838	2 Q862B6	Q862B6 botrytis ci
21	608.5	23.5	548	2 Q8SQ09	Q8SQ09 necephalico
22	604	23.3	912	2 Q7RX60	Q7RX60 neurospora
23	603.5	23.3	821	2 Q7QDS6	Q7QDS6 anopheles g
24	597.5	23.1	703	2 Q9F277	Q9F277 arabidopsis
25	593	22.9	677	2 Q9VRK9	Q9VRK9 drosophila
26	584.5	22.6	642	2 Q7QV55	Q7QV55 giardia lam
27	580.5	22.4	1394	2 Q7M6Z4	Q7M6Z4 mus musculus
28	579.5	22.4	1226	1 KFAA XENLA	Q91784 xenopus lae
29	579.5	22.4	1226	2 Q6LRM2	Q6LRM2 xenopus lae
30	579.5	22.4	1226	2 AAH70854	AAH70854 xenopus l
31	579	22.4	1463	2 Q9GYZ0	Q9GYZ0 stronglyloce

32	578.5	22.3	1058	2 Q9LZU5	Q9LZU5 arabidopsis
33	577	22.3	571	2 Q8BZ45	Q8BZ45 mus musculus
34	577	22.3	880	2 Q8BL11	Q8BL11 mus musculus
35	577	22.3	886	2 Q91WD7	Q91WD7 mus musculus
36	575.5	22.2	1231	2 Q6DINS	Q6DINS xenopus tro
37	574.5	22.2	787	2 Q6PKB2	Q6PKB2 homo sapien
38	574.5	22.2	787	2 AAH03664	AAH03664 homo sapi
39	574.5	22.2	1127	2 Q86TN3	Q86TN3 homo sapien
40	574.5	22.2	1232	1 KF4A HUMAN	Q95239 homo sapien
41	574	22.2	562	2 Q8BY99	Q8BY99 mus musculus
42	572	22.1	895	2 Q7ZUW9	Q7ZUW9 brachydanio
43	570	22.0	699	2 Q98T11	Q98T11 xenopus lae
44	569.5	22.0	915	2 Q6GR48	Q6GR48 xenopus lae
45	568.5	22.0	671	2 Q86XX7	Q86XX7 homo sapien

## ALIGNMENTS

RESULT 1  
KF22 HUMAN  
ID KF22 HUMAN STANDARD; PRT; 665 AA.  
AC Q14807; Q94814; Q9BT46;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)  
DE (Kinesin-like protein 4).  
OS Homo sapiens (Human).  
GN Name=KIF22; Synonyms=KNSL4, KID;  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]\_TaxID=9606;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96174806; PubMed=8599929;  
RA Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,  
Tsukita S., Inoue J., Yamamoto T.;  
RT "Kid, a novel kinesin-like DNA binding protein, is localized to  
RT chromosomes and the mitotic spindle."  
RL EMBO J. 15:457-467(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Lymphocytoblastoma;  
RX MEDLINE=9909123; PubMed=9790757;  
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,  
Geltinger C., Saito-Ohara F., Ikeuchi T., Matsumura M., Itakura K.,  
Kanazawa I., Sun K., Yokoyama K.K.;  
RT "Human genes for KNSL4 and MAZ are located close to one another on  
RL Chromosome 16p11.2."  
RN [3]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Brain, and Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human



QY 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQ 242  
 Db 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQ 241  
 QY 243 RSSRSHAVLVKVDORERLAPFQREGLYLIDLAGSEDRRTGNGLRKGSAINTSL 302  
 Db 242 RSSRSHAVLVKVDORERLAPFQREGLYLIDLAGSEDRRTGNGLRKGSAINTSL 301  
 QY 303 FVLGKVVADLNQGLPRVYPRYRDSKLTLLQSLGSSAHSILIANIAPERFYLTIVSALNF 362  
 Db 302 FVLGKVVADLNQGLPRVYPRYRDSKLTLLQSLGSSAHSILIANIAPERFYLTIVSALNF 361  
 QY 363 AARSKEVINPFTNESLOPHALGPVKLSQKELLGPPEAKRGPPEBEEIGSPPEMAAPAS 422  
 Db 362 AARSKEVINPFTNESLOPHALGPVKLSQKELLGPPEAKRGPPEBEEIGSPPEMAAPAS 421  
 QY 423 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERVMVLMKTVEEKDLE 482  
 Db 422 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERVMVLMKTVEEKDLE 481  
 QY 483 IERLKTQKLEAKMLAQKAEEKENHCPTM 512  
 Db 482 IERLKTQKLEAKMLAQKAEEKENHCPTM 511

RESULT 3  
 Q99LC7  
 ID Q99LC7 PRELIMINARY; PRT; 660 AA.  
 AC Q99LC7  
 DT 01-JUN-2001 (TEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
 DE Kinesin family member 22.  
 GN Names-Kif22;  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Maria M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC003427; AA03427.1; -;  
 DR HSSP; P33173; 1155.  
 DR MGD; MGI:109233; Kif22.  
 DR GO; GO:0000785; C:chromatin; IDA.  
 DR GO; GO:0005819; C:spindle; IDA.  
 DR InterPro; IPR003583; HHH\_1.

DR InterPro; IPR001752; kinesin motor.  
 DR InterPro; IPR010994; Ruva\_2\_like.  
 DR Pfam; PF00225; Kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00278; HHH1; 2.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
 KW ATP-binding; Microtubule; Motor protein.  
 SQ SEQUENCE 660 AA; 73189 MW; AA8B99477BC49B3C CRC64;

Query Match 83.4%; Score 2158.5; DB 2; Length 660;  
 Best Local Similarity 84.8%; Pred. No. 3.8e-132;  
 Matches 431; Conservative 32; Mismatches 36; Indels 9; Gaps 4;

QY 8 TQORREMAAASAAATSGAGRCRLSKIGATRRPPARVAVVLRPPVDCGTAGASDPPCV 67  
 Db 7 TQORREMAAATSGAGRCRLSKIGATRRPPARVAVVLRPPVDCGTAGASDPPCV 61  
 QY 68 RGMDSCLSEITANRNHQLKQVDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAY 127  
 Db 62 RAIDSCSLELVANWKKYQETLKQVDFAFYGEKSTQQEVYVGVQPIRLHLLLEGQNASVLAY 121  
 QY 128 GPTGAGKTHMLGSPQPGVIPPALMDLLQTLTREBGAEGRPWALSVTMTSYLEYQEKVLD 187  
 Db 122 GPTGAGKTHMLGSPQPGVIPPALMDLLQTLTREBGAEGRPWALSVTMTSYLEYQEKVLD 181  
 QY 188 LDDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQSSRS 247  
 Db 182 LDDPASGDLVIREDCRGNILIPGLTQKPISSPADFERHFLPASNRRTVGATRLNQSSRS 241  
 QY 248 HAVLLVKVDQERLAPFQREGLYLIDLAGSEDRRTGNGLRKGSAINTSLFVLGK 307  
 Db 242 HAVLLVKVDQERLAPFQREGLYLIDLAGSEDRRTGNGLRKGSAINTSLFVLGK 301  
 QY 308 VVDALNQGPRVYPRYRDSKLTLLQSLGSSAHSILIANIAPERFYLTIVSALNFARSK 367  
 Db 302 VVDALNQGPRVYPRYRDSKLTLLQSLGSSAHSILIANIAPERFYLTIVSALNFARSK 361  
 QY 368 EVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPPEBEEIGSPPEMAAPASOKL 427  
 Db 362 EVINRPTNESLOPHALGPVKLSQKELLGPPEAKRGPPEBEEIGSPPEMAAPASOKL 421  
 QY 428 SPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERVMVLMKTVEEKDLEIRLK 487  
 Db 422 SPLQKLSMDPAMLERLLSLDRLLASQSQGAPLLSTPKRERVMVLMKTVEEKDLEIRLK 481  
 QY 488 TKQKLEAKMLAQKAEEKENHCPTM 512  
 Db 482 MKQKLEAKMLAQKAEEKENHCPTM 508

RESULT 4  
 Q6P3R1  
 ID Q6P3R1 PRELIMINARY; PRT; 639 AA.  
 AC Q6P3R1  
 DT 05-JUL-2004 (TEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein MGC75575.  
 GN Name=MGC75575;  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=8364;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

	Query Match	56.5%	Score 1463;	DB 2;	Length 639;
	Best Local Similarity	60.3%;	Pred. No. 7.8e-8;		
	Matches 292;	Conservative	79;	Mismatches 103;	Indels 10;
					Gaps
Qy	30	RLSKIGATRRPPPARVAVRLRPFVDGTAGASDPCCVVRGMDSCSLEIANWRNHQETLKY	89		
Db	4	RVSILDQHKFSSARVAVRLRPFYMEKDEKAPACVVRGLDSQSLEIVNWRNQLEITMOY	63		
Qy	90	QFDAFYGERSTQODIYAGSVQPIRLHLLGGQNASVLAYGPTGAKGTHTMLGSPQPGVIP	149		
Db	64	QFDAFYGDASATQREIYMGSVCHILPHLLIGQNASVFAYGPTGAKGTHTMLGNPSPQGVIP	123		
Qy	150	RALMDLLQLTREE--GAGRPWALSVTMSYLEIYQEKVLDLDDPASGLDVLIREDCRGNIL	207		
Db	124	RAVRDLLQMTTAAGGPNENWYITITWSYEIYQEKVMDLLEPKNDLPREDKDNIL	183		
Qy	208	IPGLSOKPISFPADFERHFLPASNRRTGATLNORSRSNAVLLVKVDQRELRAPFRQR	267		
Db	184	IPGVQTKTINGPQDDEHFIPASQNTVASTKLNDRSSRSNAVLLIKVKSQSVSPFRL	243		
Qy	268	EGLKYLIDLAGSEDRNRRTGNKGLRLKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKLT	327		
Db	244	TGKLYLIDLAGSEDRNRRTGNQGIRLKESGAINSSLFTLSKVVDALNOGLPRIPYRDSKLT	303		
Qy	328	RLQDSLGSSAHSIIIANIAPRRFYLDTVSALNFAARSKVEVINPFTNESLQPHALGPV	387		
Db	304	RLIQDSLGSTASHVMIANIAPQKIFYDFLTALNFAAKSKQIINKPFOETTSIAALPA	363		
Qy	398	KLSQKELLGPPEAKRGPBEEIGSPPEMAAPASAKLSLQK--LSSMDPAMLERLL	445		
Db	364	MKRPRE----BAETAAGSRQRKSKTSDTESPNTSMDAASKRKLNALADPAVVERLL	418		
Qy	446	SLDRLLASQSGCAPLLSTPKRBMVLMKTVEEKOLEIRLTKYKEAKMLAKAE-E	504		
Db	419	KLDKILTEKGMKEAQLSTPKRERMALLKKWEESOMEIERLKEKQELQKAEIAREARLE	478		
Qy	505	KENH 508			
Db	479	KSTN 482			

RESULT 5  
AAH63896 PRELIMINARY; PRT; 639 AA.  
AC AAH63896;  
DT 25-WAR-2004 (TrEMBLrel. 27, Created)  
DT 25-WAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 25-WAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein MGC75575.  
GN MGC75575.  
OS Xenopus tropicalis, (Western clawed frog) (*Silurana tropicalis*).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus; Silurana.  
OX NCBI\_TaxID=8364;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey M., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting J., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Kzywinski M.I., Skalka U., Smalios D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RC Klein S., Strausberg R.;  
RE Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL; BC063896; AAH63896.1; --  
RW Hypothetical protein.  
SQ SEQUENCE 639 AA; F5664F9020A6A1AA CRC64;  
  
Query Match 56.5%; Score 1463; DB 2; Length 639;  
Best Local Similarity 60.3%; Pred. No. 7.8e-87;  
Matches 292; Conservative  
  
QY 30 RLSKIGATRRPPPARVAVRLRPFVDGTAGASDPCVRGMDSCSLSIANWRNHQETLK Y 89  
DB 4 RVSLDQHKKPSSARVARVAVRLRPFYKEKEBKAPAACVRLDSQSLEIVNWRNQLETMQ Y 63  
QY 90 QFDIFYGERSTQQDIYAGSVQPIRLRHLEGONASVLAYGPTGACKTHMLGSPROPVIP 149  
DB 64 QFDIFYGDSATQREIYMGVSCHILPHLLIGONASFAYGPTGACKTHMLGNPSOGVIP 123  
QY 150 RALMDLIQLTREE--GAEGRPWALSVTMSYLEIQEKLVDLLDPASGLDVIRECRGNIL 207  
DB 124 RAVRDLLQMTRTAAGPENENWTITMSVEVIQEKMDLLEPKNKDLPIREDKHNIL 183  
QY 208 IPGLSQKPDISSFADFERHFLPASNRNTVGATRLNQRSSRSNAVLLVKVDRLAPPQR 267  
DB 194 IPGVTKTINSFGDFDEHFIPASQNRTVASTKLNDRSSRSNAVLLIKVKSQQSPFKOL 243  
QY 268 EGKYLIDLAGESENRFTNGKGLRIKESGAINTSILFVLGKVVDALNOGLPRVPYRDSKIT 327  
DB 244 TKGYLIDLAGESENNRTGNGIRUKESGAINSSLFTLSKVVDALNOGLPRIPYRDSKIT 303  
QY 328 RLLODSLGSASHSLIANIAPERFYLDTVTSALMFAAKSKEVINRPFNTNESLPHALGPV 387



QY	387	VKLSQKE---	LLGPPEAKRARGPEEEIGSPPEMAAPASAKLSPLQKLSMDPAMLER	443				
DB	377	MKRPREETGHTAGSQKRRKSKNDSTE--SSPN--SSMDTAGKQKLN----	LATLDPAAVVER	429				
QY	444	LLSLDLRLAASQGSQAPLLSTPKRERMVLMKTVEEKLEIRLTKYKKELEAKMLAQKAE	503					
DB	430	LLKLDKILTEKGKKKQALLSTPKRERMALLKKWEESQMEIERLKEKQKELEQKAMEAEAR	489					
QY	504	-EKENH	508					
DB	490	LEKSNN	495					
RESULT 7								
ID	AAH70549	PRELIMINARY;	PRT;	651 AA.				
AC	AAH70549;							
DT	13-MAY-2004	(TremBLrel. 27, Created)						
DT	13-MAY-2004	(TremBLrel. 27, Last sequence update)						
DT	13-MAY-2004	(TremBLrel. 27, Last annotation update)						
DE	Hypothetical protein.							
OS	Xenopus laevis (African clawed frog).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;							
OC	Xenopodinae; Xenopus; Xenopus.							
OX	NCBI_TaxID=8355;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	TISSUE=Embryo;							
RX	MEDLINE=22341132; PubMed=12454917;							
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,							
RA	Richardson P.;							
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus							
RT	initiative.";							
RL	Dev. Dyn. 225:384-391(2002).							
RP	[2]							
RN	SEQUENCE FROM N.A.							
RC	TISSUE=Embryo;							
RX	MEDLINE=22388257; PubMed=12477932;							
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,							
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,							
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,							
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,							
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,							
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,							
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,							
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,							
RA	Bohak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,							
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,							
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,							
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,							
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,							
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,							
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,							
RA	Krzyszynski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,							
RA	Jones S.J., Marra M.A.;							
RT	"Generation and initial analysis of more than 15,000 full-length human							
RT	and mouse cDNA sequences.";							
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).							
RP	[3]							
RN	SEQUENCE FROM N.A.							
RC	TISSUE=Embryo;							
RA	Klein S., Strausberg R.;							
RL	Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.							
RL	EMBL; BC070549; AAH70549.1;							
KW	Hypothetical protein.							
QY	SEQUENCE 651 AA; 73065 MW; C2ABA561C4C53C13 CRC64;							
Query Match								
Best Local Similarity 55.5%; Score 1438; DB 2; Length 651;								
Matches 292; Conservative 80; Mismatches 100; Indels 14; Gaps								
7;								

QY	30	RLSKIGATRRPPPARVAVRLRPFVDVGTAGA--SDPCTVRGMDSCSLEIATNRNHOETLK	88	
DB	17	RVSMLDQHKKSSCARVAVRLRPYMDKDEAKATTVCVRGLDSQSLEIATNRNHOETMQ	76	
QY	89	YQFDAPGVERSTQODIYAGSVQPTLRHLLEGQNASVLAVGPTGAGKTHTMLGSPQPGVI	148	
DB	77	YQFDAPGYGDSASQREIYMGSVCHILPHLLIGQNASVAVGPTGAGKTHTMLGNPNQPGVI	136	
QY	149	PRALMDLIQLTRREGA--EGRPWALSVTMSYLIYQEKVLDLLDPASGDLVIREDCRNI	206	
DB	137	PRAVRDLLQMSRTAASAPENENWTYTINMSYVEIYQEKVMDLLEPKNKDLPFIREDKDNI	196	
QY	207	LIPGLSQKPISSPADFERHFLPASRNFYTCATNLNORSSRSHAVLLVKVDORERLAPFRQ	266	
DB	197	LIPGVYQKRWINSFADFDHFIPASQNRVTASTKLNDRSSRSHAVLLVKVKSQVVVFRQ	256	
QY	267	REGKLYLIDLAGSDNRRTNGKGLRKESGAINTSFLVGLKVVDALNQGLPRVPVYRDSKL	326	
DB	257	LTGKLYLIDLAGSDNRRTNGQGLRKESGAINSSLFTLSKVVDALNQGLPRIPYRDSKL	316	
QY	327	TRLLQDSIGGSAHSILIANIAPERFVYDTSALNFAARKEVINRPTNESIQPHALGP	366	
DB	317	TRLLQDSIGGSAHSVMTIATPEQTYFYFDLTALNFAAKSQIINKFPSETTQTVVQPA	376	
QY	387	VKLSQKE--LLGPPEAKRARGPEEEIGSPPEMAAPASAKLSPLQKLSMDPAMLER	443	
DB	377	MKRPREETGHTAGSQKRRKSKNDSTE--SSPN--SSMDTAGKQKLN----	429	
QY	444	LLSLDLRLAASQGSQAPLLSTPKRERMVLMKTVEEKLEIRLTKYKKELEAKMLAQKAE	503	
DB	430	LLKLDKILTEKGKKKQALLSTPKRERMALLKKWEESQMEIERLKEKQKELEQKAMEAEAR	489	
QY	504	-EKENH	508	
DB	490	LEKSNN	495	
RESULT 8				
ID	Q918K0	PRELIMINARY;	PRT;	651 AA.
AC	Q918K0;			
DT	01-OCT-2000	(TremBLrel. 15, Created)		
DT	01-OCT-2000	(TremBLrel. 15, Last sequence update)		
DT	01-MAR-2004	(TremBLrel. 26, Last annotation update)		
DE	Chromokinesin Xkid.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20419288; PubMed=10966104;			
RA	Funabiki H., Murray A.W.;			
RT	"The Xenopus chromokinesin Xkid is essential for metaphase chromosome			
RT	alignment and must be degraded to allow anaphase chromosome			
RT	movement.";			
RL	Cell 102:411-424(2000).			
RL	EMBL; AF267850; AAF82564.1; --			
DR	HSSP; P31373; I15S.			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0005875; C:microtubule associated complex; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0003774; F:motor activity; IEA.			
DR	InterPro; IPR010996; DNAPol_B_n_like.			
DR	InterPro; IPR003583; HHH 1.			
DR	InterPro; IPR001752; kinesin_motor.			
DR	Pfam; PF00225; Kinesin; 1.			
DR	PRINTS; P000380; KINESINHEAVY.			
DR	SMART; SM00278; HHH1; 2.			
DR	SMART; SM00129; KISC; 1.			
DR	PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.			



```
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
KW ATP-binding; Microtubule; Motor protein.  
SQ SEQUENCE 651 AA; 73091 MW; C4EF79801E603C13 CRC64;  
  
Query Match          55.5%; Score 1437; DB 2; Length 651;  
Best Local Similarity 60.1%; Pred. No. 4e-85;  
Matches 292; Conservative 80; Mismatches 100; Indels 14; Gaps 7  
  
Qy 30 RLSKIGATRRPPPARVRVAVRLRPFDVGDTAGA--SDPPCVRGMDSCSLEIANWRNRHQBTLK 148  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 17 RVSMLDQHKKSSCARVRVAVRLRPLYMDEKDAKATTCVVRGLDSQSLEIVNWRNQLETMQ 136  
  
Qy 89 YQFDAPYGRSTOODIYAGSVOPILRHLLLEGONASVLAYGPTGAKGTHMLGSGPEPGVI 148  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 77 YQFDDAFYGDSASQREIYMGVCYIIPLHLLIGNASVFAYGPTGAKGTHMLGNPNOPGVI 136  
  
Qy 149 PRALMDLLQLTREGA--EGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGI 206  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 137 PRAVRDOLLQMSRTAASAPENWTTIIMNSVEIYQEKVMOLLEPKNDLPREDKDHN 196  
  
Qy 207 LIPGLSKQTISFPADPERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQRERLAPPFQ 266  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 197 LIPGVTKQMINSFADPDEHFIPASQNRTVAETKLNDRSSRSRSHAVLLIKVQKSQQVVPFRQ 256  
  
Qy 267 REGKLYILDLAGSEDNRTGNGKLBLEKSGAINTSLFVLGVKDVALNOGLPRVPYRD SKL 326  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 257 LTGKGLYILDLAGSEDNRTGNGOIRLKESGAINSFLTSLKVVDNALNOGLPRIPYRDSL 316  
  
Qy 327 TRLLQDSLGSGSAHSILIANIAPERFYLDTVSALNFARSKVINRPFTHESLOPHALGP 386  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 317 TRLLQDSLGSGSAHSMVTMTIAPEOTYYFTUTALNFAAKSKQIINKPFSQETTQTVVQPA 376  
  
Qy 387 VKLSQKE---LLGPPPEAKRGAEPREEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLER 443  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 377 MKPREETHIAGSQKCKSKNDSTE--SSPN-SMMDTAGKQKN----LATIDPAVVER 429  
  
Qy 444 LLSLDRLLASQSQGAPLLSTPKPRRMVLMKTVEEKOILERTLTKOKELEAKMLAQKAE 503  
   : :: |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db 430 LLKDLKILTKEGKKQAQLLSTPKPRERMALLKKWEESQMEIERLKEKOLEQKAMEAEAR 489  
  
Qy 504 -EKENH 508  
   |||:  
Db 490 LEKSN 495  
  
RESULT 9  
Q7ZYL5 PRELIMINARY; PRT; 663 AA.  
ID Q7ZYLS PRELIMINARY; PRT; 663 AA.  
AC Q7ZYLS;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
De Similar to kinesin family member 22 (Fragment).  
Os Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC xenopodinae; Xenopus.  
OX NCBI_TaxID=8355;  
RN [1]  
SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RL Klein S., Strausberg R.;  
RA Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
RL EMBL; BC043733; AHA43733.1; -.  
DR HSSP; P33173; I158.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005875; C:microtubule associated complex; IEA.  
DR GO; GO:0005824; F:ATP binding; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR InterPro; IPR010996; DNaPol_B_like.  
DR InterPro; IPR003583; HH1.1.  
DR InterPro; IPR001752; kinesin motor.
```

[illegible]

Fri Nov 12 12:26:18 2004

```
DR EMBL; AJ249841; CAB71799.1; -.
DR HSSP; P33173; 115S.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; D:NAAP B_N-like.
DR InterPro; IPR003583; HH1_1.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00467; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
KW SEQUENCE 650 AA; 73027 MW; EA415BC4B9B0F4A CRC64;
Query Match 55.0%; Score 1423; DB 2; Length 650;
Best Local Similarity 59.4%; Pred. No. 3.2e-84;
Matches 287; Conservative 81; Mismatches 107; Indels 8; Gaps 4;
QY 30 RLSKIGATRRPPPARVAVRLRPFDGTAGASDP-PCVRGMDSCSLEIANWRNHQETLK 88
Db 17 RASMLDQHKKSCARVAVRLRPYRPEEKEDKVTACVRGLDSHSLEIVNWRNQLETMQ 76
QY 89 YQDFAFYGRSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAKTHMLGSPQPGVI 148
Db 77 YQDFAFYGRSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAKTHMLGSPQPGVI 136
QY 149 PRALMDLLQTR--BEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDVIREDCRNI 206
Db 137 PRVRELLQWTRMAASAPENENWTHITMSYVEIYQEKVMDLLEPKNDLPFIREDKDNI 196
QY 207 LIPGLSOKDISPAPFERHLPASNRRTVGNLRNQRSSSHAVLLVKVDORERLAPRQ 266
Db 197 LIPGLSOKDISPAPFERHLPASNRRTVGNLRNQRSSSHAVLLVKVDORERLAPRQ 256
QY 267 REGKYLIDLAGSDNRRTGNKGLRKESGAINSLFVLGKVVVDALNOGLPRVPRDLSKL 326
Db 257 LICKYLIDLAGSDNRRTGNKGLRKESGAINSLFVLGKVVVDALNOGLPRVPRDLSKL 316
QY 327 TRLLQSLGSAHSIIANIAPERRYLDTVSNALFAESKEVINRPFNTNESLOPHALGP 386
Db 317 TRLLQSLGSAHSIIANIAPERRYLDTVSNALFAESKEVINRPFNTNESLOPHALGP 376
QY 387 VKLSQKELLGPPPAKARGPEEEIGSPFMAAPASAKSLPQKLSMDPAMLERLLS 446
Db 377 MKRPREAEATTSRQRKSKTSTESSPNSSMESTGKLN---LASLDSAVVERLLK 432
QY 447 LDRLLASQSGQAPLSTPKRERVLMTVEEKDLIERIKYQKELEAKMLAQKAE-EK 505
Db 433 LDKILTEKGKKEAQLSTPKRERVLMTVEEKDLIERIKYQKELEAKMLAQKAE-EK 492
QY 506 ENH 508
Db 493 SNN 495
RESULT 11
Q6GPGO PRELIMINARY; PRT; 631 AA.
AC Q6GPGO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073177; AAH73177.1; -.
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR010994; Ruva_2_like.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00467; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Hypothetical protein; Microtubule; Motor protein.
KW SEQUENCE 631 AA; 71084 MW; 11E8C9AB86EBD3F3 CRC64;
Query Match 54.9%; Score 1422; DB 2; Length 631;
Best Local Similarity 60.6%; Pred. No. 3.6e-84;
Matches 285; Conservative 78; Mismatches 99; Indels 8; Gaps 4;
QY 43 ARRVAVRLRPFDGTAGASDP-PCVRGMDSCSLEIANWRNHQETLKYQDFAFYGRSTQ 101
Db 11 ARRVAVRLRPYRPEEKEDKVTACVRGLDSHSLEIVNWRNQLETMQYQDFAFYGRSTQ 70
QY 102 QDIYAGSVQPIRLHLLLEGONASVLAAYGPTGAKTHMLGSPQPGVTPRALMDLLQTR- 160
Db 71 REIYMGSVCHILPHLLIGONASVFAAYGPTGAKTHMLGNPDQPGVPRVRELLQWTRM 130
QY 161 -BEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDVIREDCRNIILPGLSOKPISSF 219
Db 131 AASAPENENWTHITMSYVEIYQEKVMDLLEPKNDLPFIREDKDNIILPGVTLKINSF 190
QY 220 ADPFRHFLPASNRRTVGNLRNQRSSSHAVLLVKVDORERLAPRQREGKYLIDLAGS 279
Db 191 GDFDEHFPASQNRIVASTKLNDSSRSRSHAVLLIKVQKSVQVAPRQLIGKYLIDLAGS 250
QY 280 EDNRRTGNKGLRKESGAINSLFVLGKVVVDALNOGLPRVPRDLSKLTRLLQDSLGSAH 339
Db 251 EDNRRTGNKGLRKESGAINSLFVLGKVVVDALNOGLPRVPRDLSKLTRLLQDSLGSAH 310
```

```

QY 340 SILIANIAPERRYLDTGALNFAARSKVINRPPFTNESLQPHALGPVKLSQKELGPPE 399
Db 311 SVMITNIAEQTYFDTLTALNFAAKSKQIKNPFSSRTTQTVAQPMKRPREEAETTS 370
QY 400 AXKARGPEEEIGSPMPAASAKSLPQKLSMDPAMLERLILSLDELLASQSQGA 459
Db 371 SRQKSKSDTSESSPNMESGTRKLN----LASLDSAVVERLLKDKILTEKKGKEA 426
QY 460 PLLSTPKRRMYMLTVBKEKLEIERLTKQKELEAKMLAQKAE-EKENH 508
Db 427 QLLSTPKRRMALLKKWESSQWEIERLKEKQKELEKAMEAEARLEKSN 476

RESULT 12
O60845
ID O60845 PRELIMINARY; PRT; 198 AA.
AC O60845;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Similar to kinesin-like DNA binding protein (KID).
GN Name=A-328A3.2;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9425270; PubMed=10493829;
RA Loftus B.J., Kim U.J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q."
RL Genomics 60:295-308(1999).
DR EMBL; AC002301; AAC08709.1; -.
DR HSSP; P33173; 1161.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 198 AA; 21567 MW; 1A7FDA95B5A03B37 CRC64;

Query Match 36.5%; Score 946; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 7.6e-54;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGSGTQQRREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPPVDGTAGAS 62
Db 2 AAGSGTQQRREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPPVDGTAGAS 61
QY 63 DPPCVRGMDSCSLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLEGQNA 122
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKYQFDAPYGERSTQDDIYAGSVQPIRLHLEGQNA 121
QY 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQ 182
Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQ 181
QY 183 EKV 185
Db 182 EKV 184

RESULT 13
O35232
ID O35232 PRELIMINARY; PRT; 148 AA.
AC O35232;

```

```

DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Kinesin motor protein Kif22 (Fragment).
GN Name=Kif22;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97480723; PubMed=9339368;
RA Yang Z., Hanlon D.W., Marszalek J.R., Goldstein L.S.;
RT "Identification, partial characterization, and genetic mapping of
RT kinesin-like protein genes in mouse."
RL Genomics 45:123-131(1997).
DR EMBL; AF013119; AAC39968.1; -.
DR MGD; MGI:109233; Kif22.
DR GO; GO:0000785; C:chromatin; IEA.
DR GO; GO:0005819; C:spindle; IEA.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16596 MW; 0950C265C3B63EB7 CRC64;

Query Match 27.2%; Score 704; DB 2; Length 148;
Best Local Similarity 92.6%; Pred. No. 3e-38;
Matches 137; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 133 GKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDDPA 192
Db 1 GKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDDPA 60
QY 193 SGLVIREDCRGNILIPGLSQKPSFADFEHFHPASNRVTVGATRLNQRSSRAVLL 252
Db 61 SGLVIREDCRGNILIPGLTQKPTFSFDFEQHFLPASNRVAVGATRLNQRSSRAVLL 120
QY 253 VKVDQERLAPPRQREGKLYLIDLAGSE 280
Db 121 VKVDQERLTPPRQREGKLYLIDLAGSE 148

RESULT 14
Q6L512
ID Q6L512 PRELIMINARY; PRT; 584 AA.
AC Q6L512;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative kinesin-like DNA binding protein.
GN Name=OJ1281_H05.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
RA Li Y.-E., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117265; AAT39162.1; -.
DR InterPro; IPR010996; DNaPol_B_like.

```

Fri Nov 12 12:26:18 2004

```
DR InterPro: IPR000445; HHH.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00633; HHH; 1.
DR Pfam: PF00225; Kinesin; 1.
DR PRINTS: PRO0380; KINESINHEAVY.
DR SMART: SM0129; KISC; 1.
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE: PS0067; KINESIN MOTOR DOMAIN2; 1.
DR ATP-binding; Microtubule; Motor protein.
KW
SQ SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;

Query Match      24.4%; Score 633; DB 2; Length 584;
Best Local Similarity 36.1%; Pred. No. 8.5e-33;
Matches 166; Conservative 89; Mismatches 155; Indels 50; Gaps 15;

QY 31 LSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANRNHQTILK- 88
DB 1 MATAAATQSQP---VRVILRVRLPSEANSAEAPCVGLGSHPGGEVTVQLKDQYTSRN 57

QY 89 --YQFDAYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMLGSPQPG 146
DB 58 ECKYLDATFGQESRVCIFDQESAVIPGIEGTNATVFAYGATGSGKTYTMOGTEDLPG 117

QY 147 VIPRALMDLLOLTREGEAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 206
DB 118 LIPLAVSTVALCT-----GTWC-SVEISYEVYMERCYDLEPKAREIMVLDKDGNL 170

QY 207 LIPGLSQKPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 266
DB 171 QUKGLAWVVRSLERFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223

QY 267 REGKLYLIDLAGEENRRTGNKGLRKESGAINSTLSFVLGKVDALNOGLPRVYRDSKL 326
DB 224 VKGKLNLDLAGNEDNRRTCNIGIRLOESAKINQSLFALSNIYALNKKKEPRIPYRESKL 283

QY 327 TRLQDSLGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPFNTNESLQPHALGP 386
DB 284 TRILODSLGNSHAVMIACLNPE--YQAVHTVSLAARSRHVTH-MSSASKQETPKDK 340

QY 387 VKLSQK-----ELLGPPEA-KRARGPEEEIGSPPEMAAP-----ASAKSLSPLOKL 433
DB 341 VDMEAKLRAWLESKGKTKSIQMDG-----LLSPNAIKTFLSMHKKQSGASGRVSGRGA 395

QY 434 SSMDPAMLERLL--SLDRLLAS-----QSQGAPLLSTPKR 467
DB 396 MDQDGGKIKKVLFDASAARIPAEINFHREGTQ--DIVNTTKK 433

Search completed: November 5, 2004, 18:38:41
Job time : 78.7254 secs
```

```
RESULT 15
AAT39162
ID AAT39162 PRELIMINARY; PRT; 584 AA.
AC AAT39162;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative kinesin-like DNA binding protein.
GN OUI281_H05.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Chow T.-Y., Hsing Y.-I. C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Han S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-F., Shaw J.-F.;
RA "Oryza sativa BAC OUI281_H05 genomic sequence.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 51.8496 Seconds  
(without alignments)  
3839.562 Million cell updates/sec

Title: US-10-797-893-6

Perfect score: 1768

Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTVSALNFAARSKEVIN 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	99.7	665	1 KP22 HUMAN	Q14807 homo sapien
2	1763	99.7	665	2 AAP35923	Aap35923 homo sapi
3	1762.5	88.9	660	2 Q991C7	Q991C7 mus musculu
4	1232	69.7	639	2 Q6P3R1	Q6P3R1 xenopus tro
5	1232	69.7	639	2 AAH63896	AAH63896 xenopus t
6	1202.5	68.0	651	2 Q91869	Q91869 xenopus lae
7	1202.5	68.0	651	2 AAH70549	AAH70549 xenopus l
8	1201.5	68.0	651	2 Q918K0	Q918K0 xenopus lae
9	1197.5	67.7	663	2 Q7ZYL5	Q7ZYL5 xenopus lae
10	1192.5	67.4	650	2 Q919A8	Q919A8 xenopus lae
11	1191.5	67.4	631	2 Q6GPG0	Q6GPG0 xenopus lae
12	836	47.3	198	2 Q60845	Q60845 homo sapien
13	704	39.8	148	2 Q35232	Q35232 mus musculu
14	623.5	35.3	584	2 Q6L512	Q6L512 oryza sativ
15	623.5	35.3	584	2 AAH739162	AAH739162 oryza sat
16	594.5	33.6	628	2 Q8L788	Q8L788 arabidopsis
17	594.5	33.6	664	2 Q9L288	Q9L288 arabidopsis
18	584.5	33.1	548	2 Q8SQ09	Q8SQ09 neurospora
19	578	32.7	912	2 Q7RX60	Q7RX60 neurospora
20	568.5	32.2	821	2 Q7QD56	Q7QD56 anopheles g
21	563	31.8	838	2 Q86ZB6	Q86ZB6 botrytis ci
22	559.5	31.6	784	2 Q961H5	Q961H5 drosophila
23	559	31.6	548	2 Q8N1X8	Q8N1X8 homo sapien
24	555.5	31.4	784	1 KL68 DROME	P46867 drosophila
25	552.5	31.2	898	2 Q86VS5	Q86VS5 homo sapien
26	552.5	31.2	898	2 Q8N177	Q8N177 homo sapien
27	552.5	31.2	898	2 Q7H0F3	Q7H0F3 homo sapien
28	552	31.2	1296	2 Q7RM16	Q7RM16 plasmodium
29	551.5	31.2	562	2 Q8B799	Q8B799 mus musculu
30	551.5	31.2	571	2 Q8B245	Q8B245 mus musculu
31	551.5	31.2	880	2 Q8BL11	Q8BL11 mus musculu

32 551.5 31.2 886 2 Q91WD7 Q91WD7 mus musculu  
33 551.5 31.2 895 2 Q7ZUM9 Q7ZUM9 brachydanio  
34 550 31.1 703 2 Q9FZ77 Q9FZ77 arabidopsis  
35 547 30.9 1254 2 Q94463 Q94463 dictyostell  
36 546 30.9 997 2 Q7ZXX2 Q7ZXX2 xenopus lae  
37 545.5 30.9 389 2 Q7SYZ3 Q7SYZ3 xenopus lae  
38 542.5 30.7 553 2 Q7PGH1 Q7PGH1 anopheles g  
39 541 30.6 650 2 Q6DD89 Q6DD89 xenopus lae  
40 540 30.5 750 2 Q7KSK2 Q7KSK2 drosophila  
41 540 30.5 750 2 AAS65150 AAS65150 drosophil  
42 536 30.3 784 1 KLP6\_SCHPO KLP6\_SCHPO  
43 535.5 30.3 677 2 Q9VRK9 Q9VRK9 drosophila  
44 533.5 30.2 642 2 Q7QY55 Q7QY55 giardia lam  
45 533.5 30.2 834 2 Q6PFD6 Q6PFD6 mus musculu

#### ALIGNMENTS

#### RESULT 1

KF22 HUMAN  
ID KF22\_HUMAN STANDARD; PRT; 665 AA.  
AC Q14807; Q94814; Q9BTR46;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)  
DE (Kinesin-like protein 4).  
DE Name=KIF22; Synonyms=KNSL4, KID;  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96174806; PubMed=8599929;  
RA Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,  
RA Tsukita S., Inoue J., Yamamoto T.;  
RT "Kid, a novel kinesin-like DNA binding protein, is localized to  
RT chromosomes and the mitotic spindle."  
RL EMBO J. 15:457-467(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Lymphocytoblastoma;  
MEDLINE=99009323; PubMed=9790757;  
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,  
RA Geltinger C., Saito-Onara F., Ikeuchi T., Matsumura M., Itakura K.,  
RA Kanazawa I., Sun K., Yokoyama K.K.;  
RT "Human genes for KNSL4 and MAZ are located close to one another on  
RT chromosome 16p11.2."  
RL Genomics 52:374-377(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Brain, and Lung;  
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Tohiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human



```
RESULT 3
Q99LC7 PRELIMINARY; PRT; 660 AA.
ID Q99LC7
AC Q99LC7;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Kinesin family member 22.
GN Name=Kif22;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
EMBL: BC003427; AA03427.1; -.
DR HSP; P31173; I155.
DR MGD; MGI:109233; Klf22.
DR GO; GO:000785; C:chromatin; IDA.
DR GO; GO:0005819; C:spindle; IDA.
DR InterPro; IPR03583; HHH_1.
DR InterPro; IPR001752; Kinesin motor.
DR InterPro; IPR010994; Ruva 2 like.
DR Pfam; PF00225; Kinesin; 1_.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00278; HHH1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 660 AA; 73189 MW; AA8B99477BC49B3C CRC64;

Query Match 88.9%; Score 1572.5; DB 2; Length 660;
Best Local Similarity 88.4%; Pred. No. 8.5e-114;
Matches 305; Conservative 19; Mismatches 20; Indels 1; Gaps 1;

QY 2 GRCLEKIGATRPAPRVAVRLPPFVDGTAGASDPPCVRGMDSCSLIANWRHQET 61
DB 22 GRC-VSKGGGRPPPLPVRVAVRLPPFMDGETEAKELPCVRAIDSCSLVANWKYQET 80
QY 62 LKQFQDAFYGERSTQDDIYAGSVQPIILRHLLGQNASVLAIGPTGAGKTHMLGSPQPG 121
DB 81 LKQFQDAFYGEKSTQEVYGVSVQPIILRHLLGQNASVLAIGPTGAGKTHMLGSPQPG 140
QY 122 VIPRALMDLIQLTREGAAGRPWDSVMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181

RESULT 4
Q6P3R1 PRELIMINARY; PRT; 639 AA.
ID Q6P3R1;
AC Q6P3R1;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75575.
GN Name=MGC75575;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL: BC063896; AA63896.1; -.
DR InterPro; IPR001752; Kinesin motor.
DR InterPro; IPR010994; Ruva 2 like.
DR Pfam; PF00225; Kinesin; 1_.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 69.7%; Score 1232; DB 2; Length 639;
Best Local Similarity 68.6%; Pred. No. 2.8e-87;
Matches 236; Conservative 53; Mismatches 53; Indels 2; Gaps 1;
```

Query Match 69.7%; Score 1232; DB 2; Length 639;  
 Best Local Similarity 68.6%; Pred. No. 2.8e-87;  
 Matches 236; Conservative 53; Mismatches 53; Indels 2; Gaps 1;

QY 5 RLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLSEIANWRNHQETIKY 64  
 DB 4 RVSILDOHQKPPSSARVRVAVRLRPYMEKDEKAPACVARGLDQSLSLEIVNWRNQLETMOY 63  
 QY 65 QPDAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPGVIP 124  
 DB 64 QPDAFYGDSATQREIYMGSVCHILPHLLIGONASVFAYGPTGAGKTHTMLGNPQPGVIP 123  
 QY 125 RALMDLLQLTREF--GAEGRPWALSVTMSYLEIQEKVLDDLDSPASGDLVIREDCRGNIL 182  
 DB 124 RAVRDLLQWTRTAAGPENENWYITITMSVYIEQEKVMDLLEPKNDLPREDKDNIL 183  
 QY 183 IGPLSQKPTISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVQDORERLAPRQR 242  
 DB 184 IGVQTKTINSFGDFEHPFIPASQNRVTASTKLNDRSSRSHAVLLIKVQSQVSPFROL 243  
 QY 243 EGKLYLIDLAGSEDRNRRTGNKGLRKESGAINSTLSVLGKVVDALNOGLPRVPRYRDSKLT 302  
 DB 244 TGKLYLIDLAGSEDRNRRTGNKGLRKESGAINSTLSVLGKVVDALNOGLPRVPRYRDSKLT 303  
 QY 303 RLLQDSLGSAHSILIANIAPERRFYLDVTVSALNFAARKEVIN 346  
 DB 304 RLLQDSLGSTAHSMIANIAPKQYFYFDLTALNFAAKSQIIN 347

RESULT 6  
 Q91869 PRELIMINARY; PRT; 651 AA.  
 ID Q91869  
 AC Q91869; (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
 DE Kinesin (Hypothetical protein) (Chromokinesin Xkid).  
 GN Name=kid;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20419289; PubMed=10966105;  
 RA Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,  
 RA Vernos I.;  
 RT "Xkid, a chromokinesin required for chromosome alignment on the  
 RT metaphase plate.";  
 RL Cell 102:425-435(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT

QY 5 RLSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLSEIANWRNHQETIKY 64  
 DB 4 RVSILDOHQKPPSSARVRVAVRLRPYMEKDEKAPACVARGLDQSLSLEIVNWRNQLETMOY 63  
 QY 65 QPDAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPGVIP 124  
 DB 64 QPDAFYGDSATQREIYMGSVCHILPHLLIGONASVFAYGPTGAGKTHTMLGNPQPGVIP 123  
 QY 125 RALMDLLQLTREF--GAEGRPWALSVTMSYLEIQEKVLDDLDSPASGDLVIREDCRGNIL 182  
 DB 124 RAVRDLLQWTRTAAGPENENWYITITMSVYIEQEKVMDLLEPKNDLPREDKDNIL 183  
 QY 183 IGPLSQKPTISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVQDORERLAPRQR 242  
 DB 184 IGVQTKTINSFGDFEHPFIPASQNRVTASTKLNDRSSRSHAVLLIKVQSQVSPFROL 243  
 QY 243 EGKLYLIDLAGSEDRNRRTGNKGLRKESGAINSTLSVLGKVVDALNOGLPRVPRYRDSKLT 302  
 DB 244 TGKLYLIDLAGSEDRNRRTGNKGLRKESGAINSTLSVLGKVVDALNOGLPRVPRYRDSKLT 303  
 QY 303 RLLQDSLGSAHSILIANIAPERRFYLDVTVSALNFAARKEVIN 346  
 DB 304 RLLQDSLGSTAHSMIANIAPKQYFYFDLTALNFAAKSQIIN 347

RESULT 5  
 AAH63896 PRELIMINARY; PRT; 639 AA.  
 ID AAH63896  
 AC AAH63896; (Tremblrel. 27, Created)  
 DT 25-MAR-2004 (Tremblrel. 27, Last sequence update)  
 DT 25-MAR-2004 (Tremblrel. 27, Last annotation update)  
 DE Hypothetical protein MGC75575.  
 GN MGC75575.  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus; Silurana.  
 OX NCBI\_TaxID=8364;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 RW EMBL; BC063896; AAH63896.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;



[illegible]

```

QY 5 RLSKIGATRRPPARVAVRLRPFDVGTAGA-SDDPCVGRMGDSCLSEIANWRNHQETLK 63
Db 17 RVSMLDQHKSSCARVAVRLRPFDVGTAGA-SDDPCVGRMGDSCLSEIANWRNHQETLK 76
QY 64 YQFDIFYGSRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPGVI 123
Db 77 YQFDIFYGSDASQREIYMGSVCHILLPHLLIGQNASVFAIGPTGAGKTHMLGSPQPGVI 136
QY 124 PRALMDLLQLTREEGA--EGRPWALSVMTSYLEIYQEKVLDLDDPASGLVIREDCRNI 181
Db 137 PRAVRDLQMSRTAASAPENENWYTIINNSYVEIYQEKVMDLLEPKNDLPPIREDKHNI 196
QY 182 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
Db 197 LIPGVTKMINSFADFEHFIIPASQNRRTVASTKLNDRSSRSRSHAVLLIKVKQSQVVPFRQ 256
QY 242 REGKLYLIDLAGSEDRNRGTGKGLRKESGAINTSFLVGLKVVVDALNOGLPRVPYRDSKL 301
Db 257 LTGKLYLIDLAGSEDRNRGTGKGLRKESGAINSSFTLSKVVDALNOGLPRIPYRDSKL 316
QY 302 TRLLQDSLGSAHSILITANIAPERFRFYDLTVSALNFAARKEVIN 346
Db 317 TRLLQDSLGSAHSVMTITNIAPEQTYFFDTLTALNFAAKSQIIN 361

```

```

RESULT 8
QY18K0 PRELIMINARY; PRT; 651 AA.
ID QY18K0
AC QY18K0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Chromokinesin Xkid.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20419286; PubMed=10966104;
RA Funabiki H., Murray A.W.;
RT "The Xenopus chromokinesin xkid is essential for metaphase chromosome alignment and must be degraded to allow anaphase chromosome movement."
RT Cell 102:411-424 (2000).
RL EMBL; AF267850; AAF82564.1; -.
DR HSP; P33173; I155.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; DNaPol_B_N-like.
DR InterPro; IPR003583; HH1.1.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; SM00278; HH1.2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 651 AA; 73091 MW; C4EF79801E603C13 CRC64;

```

```

Query Match 68.0%; Score 1201.5; DB 2; Length 651;
Best Local Similarity 67.8%; Pred. No. 6.8e-85;
Matches 234; Conservative 52; Mismatches 56; Indels 3; Gaps 2;
QY 5 RLSKIGATRRPPARVAVRLRPFDVGTAGA-SDDPCVGRMGDSCLSEIANWRNHQETLK 63
Db 17 RVSMLDQHKSSCARVAVRLRPFDVGTAGA-SDDPCVGRMGDSCLSEIANWRNHQETLK 76

```

```

QY 64 YQFDIFYGSRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPGVI 123
Db 77 YQFDIFYGSDASQREIYMGSVCHILLPHLLIGQNASVFAIGPTGAGKTHMLGSPQPGVI 136
QY 124 PRALMDLLQLTREEGA--EGRPWALSVMTSYLEIYQEKVLDLDDPASGLVIREDCRNI 181
Db 137 PRAVRDLQMSRTAASAPENENWYTIINNSYVEIYQEKVMDLLEPKNDLPPIREDKHNI 196
QY 182 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFRQ 241
Db 197 LIPGVTKMINSFADFEHFIIPASQNRRTVASTKLNDRSSRSRSHAVLLIKVKQSQVVPFRQ 256
QY 242 REGKLYLIDLAGSEDRNRGTGKGLRKESGAINTSFLVGLKVVVDALNOGLPRVPYRDSKL 301
Db 257 LTGKLYLIDLAGSEDRNRGTGKGLRKESGAINSSFTLSKVVDALNOGLPRIPYRDSKL 316
QY 302 TRLLQDSLGSAHSILITANIAPERFRFYDLTVSALNFAARKEVIN 346
Db 317 TRLLQDSLGSAHSVMTITNIAPEQTYFFDTLTALNFAAKSQIIN 361

```

```

RESULT 9
QY2YL5 PRELIMINARY; PRT; 663 AA.
ID QY2YL5
AC QY2YL5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to kinesin family member 22 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043733; AAF43733.1; -.
DR HSP; P33173; I155.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; DNaPol_B_N-like.
DR InterPro; IPR003583; HH1.1.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; SM00278; HH1.2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
SQ SEQUENCE 663 AA; 74538 MW; 3B98B5F3C12C9C22 CRC64;

```

```

Query Match 67.7%; Score 1197.5; DB 2; Length 663;
Best Local Similarity 67.5%; Pred. No. 1.4e-84;
Matches 233; Conservative 52; Mismatches 57; Indels 3; Gaps 2;
QY 5 RLSKIGATRRPPARVAVRLRPFDVGTAGA-SDDPCVGRMGDSCLSEIANWRNHQETLK 63
Db 30 RVSMLDQHKSSCARVAVRLRPFDVGTAGA-SDDPCVGRMGDSCLSEIANWRNHQETLK 89
QY 64 YQFDIFYGSRSTQDDIYAGSVQPIRLHLLLEGQNASVLAIGPTGAGKTHMLGSPQPGVI 123
Db 90 YQFDIFYGSDASQREIYMGSVCHILLPHLLIGQNASVFAIGPTGAGKTHMLGSPQPGVI 149
QY 124 PRALMDLLQLTREEGA--EGRPWALSVMTSYLEIYQEKVLDLDDPASGLVIREDCRNI 181

```

Db 150 PRVRELLQWTRMAASAPENENWTTITMSYVEIQKVDLLEPKNDLP IREDKDHNI 209  
 Qy 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPRQ 241  
 Db 210 LIPGVTLTKTNSFGDFEHPFIPASQNRVASTKLNDRSSRSHAVLLIKVKQSQQVAPRQ 269  
 Qy 242 REGKLYLIDLAGSBDNRRTGNKGLRKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 301  
 Db 270 LIGKLYLIDLAGSBDNRRTGNQIRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 329  
 Qy 302 TRLLQDSLGSASHSILLIANIAPERFYLDTVSALNFAARKEVIN 346  
 Db 330 TRLLQDSLGSASHSVMTNIAPETQYTFDTLTALNFAAKSQIIN 374

RESULT 10  
 Q919A8 PRELIMINARY; PRT; 650 AA.  
 ID Q919A8  
 AC Q919A8  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE Kinesin.  
 GN Name: kid;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20419289; PubMed=10966105;  
 RA Antonio C., Ferby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,  
 RA Vernos I.,  
 RA "Xkid, a chromokinesin required for chromosome alignment on the  
 RT metaphase plate.";  
 RL Cell 102:425-435 (2000).  
 DR EMBL; AJ249841; CAB71799.1; -.  
 DR HSP; P33173; 1155.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0005875; C:microtubule associated complex; IEA.  
 DR GO; GO:0005224; F:ATP binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0003774; F:motor activity; IEA.  
 DR InterPro; IPR010996; DNAPOL\_B\_N-like.  
 DR InterPro; IPR003583; HHH\_1.  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; Kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00278; HHH1; 2.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
 DR ATP-binding; Microtubule; Motor protein.  
 SQ SEQUENCE 650 AA; 73027 MW; EA8415BC4B9B0F4A CRC64;

Query Match 67.48; Score 1192.5; DB 2; Length 650;  
 Best Local Similarity 67.58; Pred. No. 3.4e-84;  
 Matches 233; Conservative 51; Mismatches 58; Indels 3; Gaps 2;  
 Qy 5 RLSKIGATRRPPPARVAVRLRPFVDGTAGADP-PCVRGMDSCSLEIANRNHOBTLK 63  
 Db 17 RASMLDQHKSSCARVAVRLRPFYMEKEEDKVPTACVGLDSSLSLEIVNWNQLETMQ 76  
 Qy 64 YQFDAPGERSQTDIYAGSVQPIRLHLEGONASVLAYPTGAGKTHTMLGSPGPGVI 123  
 Db 77 YQFDAPFYGDSQREIYMGVSVCHILPHLLIQNASVFAYGPTGAGKTHTMLGPNPQPGVI 136  
 Qy 124 PRALMDLQLQTR--EAGAGRPWALSVTMSYVEIQKVDLLEPKNDLP IREDKDHNI 191  
 Db 137 PRVRELLQWTRMAASAPENENWTTITMSYVEIQKVDLLEPKNDLP IREDKDHNI 196

Qy 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPRQ 241  
 Db 197 LIPGVTLTKTNSFGDFEHPFIPASQNRVASTKLNDRSSRSHAVLLIKVKQSQQVAPRQ 256  
 Qy 242 REGKLYLIDLAGSBDNRRTGNKGLRKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 301  
 Db 257 LIGKLYLIDLAGSBDNRRTGNQIRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 316  
 Qy 302 TRLLQDSLGSASHSILLIANIAPERFYLDTVSALNFAARKEVIN 346  
 Db 317 TRLLQDSLGSASHSVMTNIAPETQYTFDTLTALNFAAKSQIIN 361

RESULT 11  
 Q6GPG0 PRELIMINARY; PRT; 631 AA.  
 ID Q6GPG0  
 AC Q6GPG0  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Kzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC073177; AAH73177.1; -.  
 DR InterPro; IPR01752; kinesin\_motor.  
 DR InterPro; IPR010994; RuVA\_2\_like.  
 DR Pfam; PF00225; Kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
 KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.

SQ SEQUENCE 631 AA; 71084 MW; 11B8C9AB86EBD3F3 CRC64;  
Query Match 67.4%; Score 1191.5; DB 2; Length 631;  
Best Local Similarity 69.6%; Pred. No. 3.9e-84;  
Matches 231; Conservative 48; Mismatches 50; Indels 3; Gaps 2;  
QY 18 ARVRVAVRLRPVVDGTAGASDP-PCVRGMDSCSLFIANWRHQTLLKYQDAFYGERSTQ 76  
Db 11 ARVRVAVRLRPVVEEKDKVPTACVRGLDHSLSLEIVNWRNQLTMOYQDAFYGDSASQ 70  
QY 77 QDIYAGSVQPIRLHLEGONASVLAGPTGAGKTHMLGSPQPGVPIPRALMDLLQLTR- 135  
Db 71 REIYMGSVCHILPHLLIGONASVFAVPTGAGKTHMLGNPDQPGVIFRAVRELLQMRM 130  
QY 136 -BEGAGRGWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSP 194  
Db 131 ASAPENENWTYITMSYVEIYQEKVMOLLEPKNDLPREDKDNILIPGVTLKINSF 190  
QY 195 ADPERHFLPASNRVTGATRLNORSRSHAVLLVKVDQERLAPFRQEGKLYLIDLGS 254  
Db 191 GDFDEHIFPASQNRITVASTKLANDRSRSHAVLLIKVQSQVAPPRQIGKLYLIDLGS 250  
QY 255 EDNRRTGKGLRKESGATNTSLFVLGKVVDALNOGLPRVPYRDSKLTRELQDSLGSAAH 314  
Db 251 EDNRRTGNGIRUKESGAINSSFTLSKVVDALNOGLPRIPYRDSKLTRELQDSLGSAAH 310  
QY 315 SILIANIAPERFYLDTVSALNPAARKEVIN 346  
Db 311 SVMITNIAPEQTYFTLTALNPAARSKQIIN 342  
RESULT 12  
O60845 PRELIMINARY; PRT; 198 AA.  
AC O60845  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Similar to kinesin-like DNA binding protein (KID).  
GN Name=A-328A3.2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
RP MEDLINE=94425270; PubMed=10493829;  
RA Loftus B.J., Kim U.J., Sneddon V.P., Kalush F., Brandon R.,  
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,  
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,  
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
RT "Genome duplications and other features in 12 Mb of DNA sequence from  
human chromosome 16p and 16q."  
RL Genomics 60:295-308(1999).  
DR EMBL; AC002301; AAC08709.1; -.  
DR HSP; P33173.1; I6I.  
DR GO; GO:0005875; C:microtubule associated complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR InterPro; IPR001752; Kinesin\_motor.  
DR Pfam; PF00225; kinesin; 1.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
KW ATP-binding; Microtubule; Motor protein.  
SQ SEQUENCE 198 AA; 21567 MW; 1A7FDA95B5A03B37 CRC64;  
Query Match 47.3%; Score 836; DB 2; Length 198;  
Best Local Similarity 100.0%; Pred. No. 4.2e-57;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVAVRLRPVVDGTAGASDP-PCVRGMDSCSLFIANWRHQT 61  
Db 26 GRCRLSKIGATRRPPPARVAVRLRPVVDGTAGASDP-PCVRGMDSCSLFIANWRHQT 85

QY 62 LKQDFAPYGERSTQODIYAGSVQPIRLHLEGONASVLAGPTGAGKTHMLGSPQPG 121  
Db 86 LKQDFAPYGERSTQODIYAGSVQPIRLHLEGONASVLAGPTGAGKTHMLGSPQPG 145  
QY 122 VIPRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKV 160  
Db 146 VIPRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKV 184  
RESULT 13  
O35232 PRELIMINARY; PRT; 148 AA.  
AC O35232  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Kinesin motor protein KIF22 (Fragment).  
GN Name=Kif22;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
RP MEDLINE=97480723; PubMed=9339368;  
RA Yang Z., Hanlon D.W., Marszalek J.R., Goldstein L.S.;  
RT "Identification, partial characterization, and genetic mapping of  
kinesin-like protein genes in mouse."  
RL Genomics 45:123-131(1997).  
DR EMBL; AF013119; AAC39968.1; -.  
DR MGD; MGI:109233; Kif22.  
DR GO; GO:0000785; C:chromatin; IEA.  
DR GO; GO:0005819; C:spindle; IEA.  
DR InterPro; IPR001752; kinesin\_motor.  
DR Pfam; PF00225; kinesin; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
DR PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.  
KW ATP-binding; Microtubule; Motor protein.  
FT NON\_TER 1  
FT NON\_TER 148  
FT NON\_TER 148  
SQ SEQUENCE 148 AA; 16596 MW; 0950C265C3B63EB7 CRC64;  
Query Match 39.8%; Score 704; DB 2; Length 148;  
Best Local Similarity 92.6%; Pred. No. 5.5e-47;  
Matches 137; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
QY 108 GKTHMLGSPQPGVPIPRALMDLLQLTRREGAGRPWALSVTMSYLEIYQEKVLDLDP 167  
Db 1 GKTHMLGSPQPGVPIPRALMDLLQLARESAEGRPWDVSVMSYLEIYQEKVLDLDP 60  
QY 168 SGDAVIREDCRGNILIPGLSQPISSFADFERHFLPASNRVTGATRLNORSRSHAVLL 227  
Db 61 SGDLVIREDCRGNILIPGLTQKPIITSFDSDFQHFLLPASNRVAVGATRLNORSRSHAVLL 120  
QY 228 VKVDQERLAPFRQEGKLYLIDLAGE 255  
Db 121 VKVDQERLTPFRQEGKLYLIDLAGE 148  
RESULT 14  
Q6L512 PRELIMINARY; PRT; 584 AA.  
AC Q6L512  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Putative kinesin-like DNA binding protein.  
GN Name=CJ1281.H05.2;  
OS Oryza sativa (Japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
 RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,  
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,  
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.,  
 RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AC117265; AAT39162.1; -  
 DR InterPro; IPR010996; DNaPol\_B\_Nlike.  
 DR InterPro; IPR000445; HHh.  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00633; HHH; 1.  
 DR Pfam; PF00225; Kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS0067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW ATP-binding; Microtubule; Motor protein.  
 SQ SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;  
 Query Match 35.3%; Score 623.5; DB 2; Length 584;  
 Best Local Similarity 40.9%; Pred. No. 6.4e-40;  
 Matches 141; Conservative 68; Mismatches 113; Indels 23; Gaps 7;  
 QY 6 LSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLK- 63  
 Db 1 MATAAATQSQP---VRVVLVRPHLPSEANSAEAPCVGLLGHGPGVTVOLKQYTSRN 57  
 QY 64 --YQFDAPYGRSTQDDIYAGSVQPIILRHLEGGONASVLAAYGPTGAGKTHMLGSPQPG 121  
 Db 58 ECKYKLDAPFGQESRVCEIFDQEVSAVIGPIEGTNATVFAYGATGSGKTYTMOGTEDLPG 117  
 QY 122 VIPALMDLLQLTREGAEGRPWALSVTMSVLEIYQEKVLDLDPASGDIVIREDCRNI 181  
 Db 118 LIPLAVSTVLALCT-----GTWC-SVEISYIYVMERCYDLLEPKAREIMVLDDKQGNL 170  
 QY 182 LIPGLSKPISSFADFERRHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQRLAPPRQ 241  
 Db 171 QLKGLAWPVRSLEEFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301  
 Db 224 VKGKLNILDLAGNEDNRRTCNEGIRLQESAKINQSLFALSNNVISALNKKKEPRIPYRESKL 283  
 QY 302 TRILQDSLGSASHSILIANIANTAPERFYLDITVSALNFAARKEVIN 346  
 Db 284 TRILQDSLGSASHAVMTACLNPVE--YQEAHVTVSLAARSHVTVN 326  
 RESULT 15  
 AAT39162 PRELIMINARY; PRT; 584 AA.  
 ID AAT39162;  
 AC AAT39162;  
 DT 01-JUN-2004 (TrEMBLrel. 27, Created)  
 DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Putative kinesin-like DNA binding protein.  
 GN OJ1281.H05.2.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,

RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,  
 RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,  
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.,  
 RA "Oryza sativa BAC OJ1281.H05 genomic sequence";  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC117265; AAT39162.1; -  
 SQ SEQUENCE 584 AA; 64124 MW; EF829C3A15F9190F CRC64;  
 Query Match 35.3%; Score 623.5; DB 2; Length 584;  
 Best Local Similarity 40.9%; Pred. No. 6.4e-40;  
 Matches 141; Conservative 68; Mismatches 113; Indels 23; Gaps 7;  
 QY 6 LSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLK- 63  
 Db 1 MATAAATQSQP---VRVVLVRPHLPSEANSAEAPCVGLLGHGPGVTVOLKQYTSRN 57  
 QY 64 --YQFDAPYGRSTQDDIYAGSVQPIILRHLEGGONASVLAAYGPTGAGKTHMLGSPQPG 121  
 Db 58 ECKYKLDAPFGQESRVCEIFDQEVSAVIGPIEGTNATVFAYGATGSGKTYTMOGTEDLPG 117  
 QY 122 VIPALMDLLQLTREGAEGRPWALSVTMSVLEIYQEKVLDLDPASGDIVIREDCRNI 181  
 Db 118 LIPLAVSTVLALCT-----GTWC-SVEISYIYVMERCYDLLEPKAREIMVLDDKQGNL 170  
 QY 182 LIPGLSKPISSFADFERRHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQRLAPPRQ 241  
 Db 171 QLKGLAWPVRSLEEFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301  
 Db 224 VKGKLNILDLAGNEDNRRTCNEGIRLQESAKINQSLFALSNNVISALNKKKEPRIPYRESKL 283  
 QY 302 TRILQDSLGSASHSILIANIANTAPERFYLDITVSALNFAARKEVIN 346  
 Db 284 TRILQDSLGSASHAVMTACLNPVE--YQEAHVTVSLAARSHVTVN 326  
 Search completed: November 5, 2004, 18:38:42  
 Job time : 52.8496 secs

This page blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:11 ; Search time 72.979 Seconds  
(without alignments)  
3839.562 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKMLAQAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2467	99.8	665	1 KF22 HUMAN	Q14807 mus sapien
2	2467	99.8	665	2 AAP35923	Aap35923 homo sapi
3	2129.5	86.1	660	2 Q91LC7	Q91LC7 mus musculu
4	1463	59.2	639	2 Q6P3R1	Q6P3R1 xenopus tro
5	1463	59.2	639	2 AAH63896	AAH63896 xenopus t
6	1438	58.2	651	2 Q91869	Q91869 xenopus lae
7	1438	58.2	651	2 AAH70549	AAH70549 xenopus l
8	1437	58.1	651	2 Q918K0	Q918K0 xenopus lae
9	1428	57.8	663	2 Q7ZYL5	Q7ZYL5 xenopus lae
10	1423	57.6	650	2 Q919A8	Q919A8 xenopus lae
11	1422	57.5	631	2 Q6GP00	Q6GP00 xenopus lae
12	836	33.8	198	2 Q60845	Q60845 mus sapien
13	704	28.5	148	2 Q32322	Q32322 mus musculu
14	633	25.6	584	2 Q6L512	Q6L512 oryza sativ
15	633	25.6	584	2 AAT39162	AAT39162 oryza sat
16	616	24.9	628	2 Q961H5	Q961H5 drosophila
17	614	24.8	628	2 Q8L7B8	Q8L7B8 arabidopsis
18	614	24.8	664	2 Q9L288	Q9L288 arabidopsis
19	612	24.8	784	1 KL68 DROME	P46867 drosophila
20	611.5	24.7	838	2 Q86ZB6	Q86ZB6 botrytis ci
21	608.5	24.6	548	2 Q8SQ09	Q8SQ09 neophalet
22	604	24.4	912	2 Q7RX60	Q7RX60 encephalit
23	603.5	24.4	821	2 Q7QDS6	Q7QDS6 anopheles g
24	597.5	24.2	703	2 Q9FZ77	Q9FZ77 arabidopsis
25	593	24.0	677	2 Q9VRK9	Q9VRK9 drosophila
26	594.5	23.6	642	2 Q7QY55	Q7QY55 giardia lam
27	590.5	23.5	1394	2 Q7M624	Q7M624 mus musculu
28	579.5	23.4	1226	1 KF4A XENIA	Q91784 xenopus lae
29	579.5	23.4	1226	2 Q6IRM2	Q6IRM2 xenopus lae
30	579.5	23.4	1226	2 AAH70854	AAH70854 xenopus l
31	579	23.4	1463	2 Q9GY20	Q9GY20 strongyloce

Q8BZ45 mus musculu  
Q8B111 mus musculu  
Q91WD7 mus musculu  
Q6A1N5 xenopus tro  
Q6PKB2 homo sapien  
AAH03664 homo sapi  
Q8B1N3 homo sapien  
Q8B239 homo sapien  
Q8BY99 mus musculu  
Q7ZUW9 brachydanio  
Q9B111 xenopus lae  
Q6GR48 xenopus lae  
Q85XX7 homo sapien  
Q86VS5 homo sapien

#### ALIGNMENTS

#### RESULT 1

KF22 HUMAN  
ID KF22 HUMAN STANDARD; PRT; 665 AA.  
AC Q14807; Q94814; Q9B746;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein)  
DE (Kinesin-like protein 4).  
GN Name=KIF22; Synonyms=KNSL4, KID;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96174806; PubMed=8599929;  
RA Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S.,  
RA Teukita S., Inoue J., Yamamoto T.;  
RT "Kid, a novel kinesin-like DNA binding protein, is localized to  
RT chromosomes and the mitotic spindle.";  
RL EMO J. 15:457-467(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Lymphocytoblastoma;  
RX MEDLINE=99009323; PubMed=9790757;  
RA Song J., Murakami H., Yang Z.Q., Koga C., Adati N., Murata T.,  
RA Geltinger C., Saito-Ohara F., Ikeuchi T., Matsumura M., Itakura K.,  
RA Kanazawa I., Sun K., Yokoyama K.K.;  
RT "Human genes for KNSL4 and MAZ are located close to one another on  
RL chromosome 16p11.2.";  
RN [3]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Brain, and Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uadin T.B., Toohyuk S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human

RT	and mouse cDNA sequences. #;	62	LKQFDAPYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG	121
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	86	LKQFDAPYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG	145
RN	[4]			
RP	INTERACTION WITH KIAH1, AND DEGRADATION.			
RX	PubMed=1146551; DOI=10.1038/sj.onc.1204002;			
RY	Germani A., Bruzzoni-Giovanelli H., Fellous A., Gisselbrecht S.,			
RA	Varin-Blank N., Calvo F.;			
RB	"KIAH-1 interacts with alpha-tubulin and degrades the kinesin kid by			
RC	the proteasome pathway during mitosis. #;			
RD	Oncogene 19:5997-6006(2000).			
RE	that is involved in spindle formation and			
RF	the movements of chromosomes during mitosis and meiosis. Binds to			
RG	microtubules and to DNA.			
RH	-1- SUBCELLULAR LOCATION: Nuclear.			
RI	-1- PTM: Ubiquitinated; mediated by KIAH1 and leading to its			
RJ	subsequent proteasomal degradation (Probable).			
RK	-1- SIMILARITY: Belongs to the kinesin-like protein family.			
RL	-----			
RM	This SWISS-PROT entry is copyright. It is produced through a collaboration			
RN	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
RO	the European Bioinformatics Institute. There are no restrictions on its			
RP	use by non-profit institutions as long as its content is in no way			
RQ	modified and this statement is not removed. Usage by and for commercial			
RR	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
RS	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
RT	-----			
RU	EMBL; AB017430; BAA33019.2; -			
RV	EMBL; AB017435; BAA33063.1; -			
RW	EMBL; AB017333; BAA33063.1; JOINED.			
RX	EMBL; AB017334; BAA33063.1; JOINED.			
RY	EMBL; BC004352; AAH04352.1; -			
RZ	EMBL; BC028155; AAH28155.1; -			
SA	HSP; P33173; I155.			
SB	Genew; HGNC:6391; KIF22.			
SC	MTM; 603213; -			
SD	GO; GO:0000776; C:kinetochore; TAS.			
SE	GO; GO:0005634; C:nucleus; TAS.			
SF	GO; GO:0003677; F:DNA binding; TAS.			
SG	GO; GO:0003777; F:microtubule motor activity; TAS.			
SH	GO; GO:0007067; P:mitosis; TAS.			
SI	InterPro; IPR003583; HHH 1.			
SJ	InterPro; IPR001752; kinesin motor.			
SK	InterPro; IPR010994; Ruva_2-like.			
SL	Pfam; PF00225; Kinesin; 1.			
SM	PRINTS; PR00380; KINESINHEAVY.			
SN	SMART; SM00278; Hhh1; 2.			
SO	SMART; SM00129; KISC; 1.			
SP	PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.			
SQ	PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.			
SR	ATP-binding; Coiled coil; DNA-binding; Microtubule; Motor protein;			
SS	Nuclear protein; Ubl conjugation.			
ST	DOMAIN 40 299			
SV	Kinesin-motor.			
SW	Coiled coil (Potential).			
SX	ATP (Potential).			
SY	Missing (in Ref. 2).			
SZ	S -> KV (in Ref. 2).			
TA	HTMGSEQGVIPRALMDLLQLTREAGRPWA -> TH			
TB	AGQPRATGDPAGSHGPPAAHKGCGCRPANG (in Ref.			
TC	2).			
TD	V -> A (in Ref. 2).			
TE	APASAKLSPLQKSSMDPAMLERLLSLDLRLASQSG			
TF	-> SSSLCPLPQPTPEAKAATWPCGAPQLGFSACLPG			
TG	P (in Ref. 2).			
TH	ENHCFPMR -> RTIIVPQSG (in Ref. 2).			
TI	SEQUENCE 665 AA; 73262 MW; C6C0AC96741DD387 CRC64;			
TJ	Query Match			
TK	Best Local Similarity			
TL	Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
TM	99.8%; Score 2467; DB 1; Length 665;			
TN	Pred. No. 4.3e-153;			
TO	2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61			
TP	26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85			
TQ	62 LKQFDAPYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121			
TR	86 LKQFDAPYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 145			
TS	122 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 181			
TT	146 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 205			
TU	Query Match			
TV	Best Local Similarity			
TV	Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
TV	99.8%; Score 2467; DB 1; Length 665;			
TV	Pred. No. 4.3e-153;			
TV	2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61			
TV	26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85			
TV	62 LKQFDAPYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121			
TV	86 LKQFDAPYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 145			
TV	122 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 181			
TV	146 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRGN 205			



```
QY 182 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPRQ 241
Db 206 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPRQ 265
QY 242 REGKYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVYRDSKL 301
Db 266 REGKYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVYRDSKL 325
QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNLSLOPHALGP 361
Db 326 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNLSLOPHALGP 385
QY 362 VKLSQKELLGPPPEAKRARGPPEEIBGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 421
Db 386 VKLSQKELLGPPPEAKRARGPPEEIBGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 445
QY 422 LDRLLASQSGCAPLLSTPKRERVMVKMTVEEKDLEIRLTKOKELKAKMLAKABEKE 481
Db 446 LDRLLASQSGCAPLLSTPKRERVMVKMTVEEKDLEIRLTKOKELKAKMLAKABEKE 505
QY 482 NHCPTM 487
Db 506 NHCPTM 511

RESULT 3
Q99LC7
ID Q99LC7 PRELIMINARY; PRT; 660 AA.
AC Q99LC7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin family member 22.
GN Name=Kif22;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Kzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
EMBL: BC003427; AAH03427.1; -.
DR HSP; P31173; 1155.
DR MGD; MGI:109233; Kif22.
DR GO; GO:0000785; C:chromatin; IDA.
DR GO; GO:0005819; C:spindle; IDA.
DR InterPro; IPR003583; HHH_1.
```

```
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR010994; Ruva_2_like.
DR Pfam; PF00225; Kinesin_1_
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00278; HHH1; 2.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 660 AA; 73189 MW; AA8B99477BC49B3C CRC64;

Query Match 86.1%; Score 2129.5; DB 2; Length 660;
Best Local Similarity 86.3%; Pred. No. 5.5e-131;
Matches 422; Conservative 29; Mismatches 33; Indels 5; Gaps 3;

QY 2 GRCLSLKIGATRRPPPARVAVRLRPVVDGTAGASDPPCVRGWDSCLSETANRNHOET 61
Db 22 GRC-VSGKGLRGRRPPLARVAVRLRPFMDGETEAKELPCVRAIDSCSELVANWKYQET 80
QY 62 LKQFDIFYGERSTQQDIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHTMLGSPQPG 121
Db 81 LKQFDIFYGEKSTQQEVVGVQPIRLHLLLEGQNASVLAYGPTGAGKTHTMLGSPQPG 140
QY 122 VIPRALMDLLQLTRECGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181
Db 141 VIPRALMDLLQLAAREESAEGRPMDVSVAMSYLEYQEKVLDLDPASGDLVIREDCRNI 200
QY 182 LIPCLSKPISSEFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPRQ 241
Db 201 LIPGLTQKPIITSPDFEQHFPLASNRRAVATRLNQRSSRSHAVLLVKVDQERLTPPRQ 260
QY 242 REGKYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVYRDSKL 301
Db 261 REGKYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVYRDSKL 320
QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNLSLOPHALGP 361
Db 321 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNLSLOPHALGP 380
QY 362 VKLSQKELLGPPPEAKRARGPPEEIBGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 421
Db 381 VKLSQKELLGPPPEAKRARGPPEEIBGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 440
QY 422 LDRLLASQSGCAPLLSTPKRERVMVKMTVEEKDLEIRLTKOKELKAKMLAKABEKE 478
Db 441 MERLLGSGSQGTPLNTPKRERVMVKMTVEEKDLEIRLTKOKELKAKMLAKABEKE 500
QY 479 EKENHCPMT 487
Db 501 EKEN-TPTI 508

RESULT 4
Q6P3R1
ID Q6P3R1 PRELIMINARY; PRT; 639 AA.
AC Q6P3R1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75575.
GN Name=MGC75575;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RA Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC063896; AAH63896.1; -  
DR InterPro; IPR001752; Kinesin motor.  
DR Pfam; PF00225; Kinesin; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS04011; KINESIN MOTOR DOMAIN1; 1.  
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.  
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.  
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 59.2%; Score 1463; DB 2; Length 639;  
Best Local Similarity 60.3%; Pred. No. 2.4e-87;  
Matches 292; Conservative 79; Mismatches 103; Indels 10; Gaps 4;

QY 5 RLSKIGATRRPPPARVAVRLRPFVDGTAGADPPCVRGMDSCSLSIANWRHQTLYK 64  
DB 4 RVSLDQHKPPSARVAVRLRPFYMEKEDEKAPACVRLGDSQSLEIWNRLQETMQY 63

QY 65 QFDIFYGERSTQODIYAGSVQPIRLHLEGONASVLAAYGTGAGKTHMLGSPQGVIP 124  
DB 64 QFDIFYGDSATQREIYMGSVCHILPHLLIGONASVFAFGTGTGAGKTHMLGNPSQGVIP 123

QY 125 RALMDLLQLTREE--GAEGRPWALSVMYSLEYQEKVLDLDPASGDLVIREDCRGNIL 182  
DB 124 RAVRDLLQMTTAAAGPENENWTITMSYVEIYQEKVMDLLEPKKNDLPPIREDKDHNIL 183

QY 183 IGPLSQKPISSPADFERHFLPASNRVTGATRLNQRSSRHAVLLVKVDQERLAPRQR 242  
DB 184 IPGVTQKTINSFGDFDEHFI PASQNRVTASTKLNDSSRHAVLLIKVQSQVSPFRL 243

QY 243 EGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVKVDALNQLPRVYRDSKLT 302  
DB 244 TGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVKVDALNQLPRVYRDSKLT 303

QY 303 RLLQDSLGSASHILIANIAPERFDLTVALNFAARKEVINRPTNLSLQHALGPV 362  
DB 304 RLLQDSLGSASHILIANIAPERFDLTVALNFAARKEVINRPTNLSLQHALGPV 363

QY 363 KLSQKELLGPENKARPEEIEGSPPEMAAPASAKLSPLQK--LSSMDPAMLERLL 420  
DB 364 MKRPRE-----EAETAAGSRQKSKTSTSTESSNTSMDAASKRLNLAALDPVAVRELL 418

QY 421 SLDRLLASQSGOAPLLSTPKRRMVLMTVEEKDLIERLTKTKQKELEAKQAKAE-E 479  
DB 419 KLDKILTEGKKEAQLLSTPKRRMVLMTVEEKDLIERLTKTKQKELEAKQAEAEARLE 478

QY 480 KENH 483  
DB 479 KSTN 482

RESULT 5  
AAH63896 PRELIMINARY; PRT; 639 AA.  
AC AAH63896;  
DT 25-MAR-2004 (T-EMBLrel. 27, Created)  
DT 25-MAR-2004 (T-EMBLrel. 27, Last sequence update)  
DT 25-MAR-2004 (T-EMBLrel. 27, Last annotation update)  
DE Hypothetical protein MGC75575.  
GN MGC75575.  
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodidae; Xenopus; Silurana.  
OX NCBI\_TaxID=8364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
EX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RA Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC063896; AAH63896.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 639 AA; 71390 MW; F5664F9020A6A1AA CRC64;

Query Match 59.2%; Score 1463; DB 2; Length 639;  
Best Local Similarity 60.3%; Pred. No. 2.4e-87;  
Matches 292; Conservative 79; Mismatches 103; Indels 10; Gaps 4;

QY 5 RLSKIGATRRPPPARVAVRLRPFVDGTAGADPPCVRGMDSCSLSIANWRHQTLYK 64  
DB 4 RVSLDQHKPPSARVAVRLRPFYMEKEDEKAPACVRLGDSQSLEIWNRLQETMQY 63

QY 65 QFDIFYGERSTQODIYAGSVQPIRLHLEGONASVLAAYGTGAGKTHMLGSPQGVIP 124  
DB 64 QFDIFYGDSATQREIYMGSVCHILPHLLIGONASVFAFGTGTGAGKTHMLGNPSQGVIP 123

QY 125 RALMDLLQLTREE--GAEGRPWALSVMYSLEYQEKVLDLDPASGDLVIREDCRGNIL 182  
DB 124 RAVRDLLQMTTAAAGPENENWTITMSYVEIYQEKVMDLLEPKKNDLPPIREDKDHNIL 183

QY 183 IGPLSQKPISSPADFERHFLPASNRVTGATRLNQRSSRHAVLLVKVDQERLAPRQR 242  
DB 184 IPGVTQKTINSFGDFDEHFI PASQNRVTASTKLNDSSRHAVLLIKVQSQVSPFRL 243

QY 243 EGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVKVDALNQLPRVYRDSKLT 302  
DB 244 TGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGVKVDALNQLPRVYRDSKLT 303

QY 303 RLLQDSLGSASHILIANIAPERFDLTVALNFAARKEVINRPTNLSLQHALGPV 362  
DB 304 RLLQDSLGSASHILIANIAPERFDLTVALNFAARKEVINRPTNLSLQHALGPV 363

QY 363 KLSQKELLGPENKARPEEIEGSPPEMAAPASAKLSPLQK--LSSMDPAMLERLL 420  
DB 364 MKRPRE-----EAETAAGSRQKSKTSTSTESSNTSMDAASKRLNLAALDPVAVRELL 418

QY 421 SLDRLLASQSGOAPLLSTPKRRMVLMTVEEKDLIERLTKTKQKELEAKQAKAE-E 479  
DB 419 KLDKILTEGKKEAQLLSTPKRRMVLMTVEEKDLIERLTKTKQKELEAKQAEAEARLE 478

QY 480 KENH 483  
DB 479 KSTN 482



QY	362	VKL	SOKE	---	LLGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLER	418				
DB	377	MKR	PREETHIAGSQKRSKNDST	---	SSPN-SSMDTAGKQKLN---	LATLDPVAVR	429			
QY	419	LLSLDRLLASQSQGAPLLSTPKRERMVMTVEKLEIERLKTKYKOLEAKMAQKAE	478							
DB	430	LLKLDKILTEKGGKKAQLLSTPKRERMVMTVEKLEIERLKTKYKOLEAKMAQKAE	489							
QY	479	-EKENH	483							
DB	490	LEKSN	495							
RESULT 7										
AAH70549	PRELIMINARY; PRT; 651 AA.									
AC	AAH70549;									
DT	13-MAY-2004	(TREMBlrel. 27, Created)								
DT	13-MAY-2004	(TREMBlrel. 27, Last sequence update)								
DE	Hypothetical protein.									
OS	Xenopus laevis (African clawed frog).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;									
OC	Xenopodinae; Xenopus; Xenopus.									
OX	NCBI_TaxID=8355;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE=Embryo;									
RX	MEDLINE=22341132; PubMed=12454917;									
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,									
RA	Richardson P.;									
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus									
RT	initiative".									
RL	Dev. Dyn. 225:384-391(2002).									
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE=Embryo;									
RX	MEDLINE=22388257; PubMed=12477932;									
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,									
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,									
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,									
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,									
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,									
RA	Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,									
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,									
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,									
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,									
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,									
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,									
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,									
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,									
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,									
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,									
RA	Krzywinski M.I., Skalska U., Smalusi D.E., Schnerch A., Schein J.E.,									
RA	Jones S.J., Marra M.A.;									
RT	"Generation and initial analysis of more than 15,000 full-length human									
RT	and mouse cDNA sequences.";									
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).									
RP	SEQUENCE FROM N.A.									
RN	TISSUE=Embryo;									
RC	TISSUE=Embryo;									
RA	Klein S., Strausberg R.;									
RL	Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.									
DR	EMBL; BC070549; AAH70549.1; --									
KW	Hypothetical protein.									
SQ	SEQUENCE 651 AA; 73065 MW; CZABA561C4C53C13 CRC64;									
Query Match										
Best Local Similarity 58.2%; Score 1438; DB 2; Length 651;										
Matches 292; Conservative 80; Mismatches 100; Indels 14; Gaps 7;										

QY	5	RLSIGATRRPPPARVAVRLRPFDVGTAGA--SDPPCVRGMDSCSLSEIANWRNHQETLK	63				
DB	17	RVSMLDQHKSSCARVAVRLRPYMDKDEAKATTVVCRGLDSQSLEIWNWNOLETMQ	76				
QY	64	YQDAFYGERSTQDIYAGSVQPIRLHLLRGONASVLAAYGPTGAGKTHMLGSPQPGVI	123				
DB	77	YQDAFYGDASQREIYMGSVCHILPHLLIGONASVFAYGPTGAGKTHMLGNPNQPGVI	136				
QY	124	PRALMDLLQUTREGA--EGRPWALSVTMVSLEYIQEKVLDLLDPASGDLVIREDCGNI	181				
DB	137	PRAVRDLQMSRTAASAPENENWTYTTINMSVETIQEKVMDLLEPKNDLPREDKHNI	196				
QY	182	LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQERLAPFRQ	241				
DB	197	LIPGVTOQMINSFADFDEHFPASQNRITVASTKLNDRSSSHAVLLIKVQSQQVVPFRQ	256				
QY	242	REGKLYLIDLAGSEDRNRRTGNKGLRKESGAINSTLFLVGLKVVDALNOGLPRVYRDSKL	301				
DB	257	LTGKLYLIDLAGSEDRNRRTGNQIRLKESGAINSTLFLSKVVDALNOGLPRIYRDSKL	316				
QY	302	TRLQDSLGSSAHSILITANIPERRFYLDTVSALNEARSKVEVINRPTNESLOPHALGP	361				
DB	317	TRLQDSLGSSAHSVMTITIAPEQTYFTLTALNFAAKSQIINKPFSQETTQTVVQPA	376				
QY	362	VKL	SOKE	---	LLGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLER	418	
DB	377	MKR	PREETHIAGSQKRSKNDST	---	SSPN-SSMDTAGKQKLN---	LATLDPVAVR	429
QY	419	LLSLDRLLASQSQGAPLLSTPKRERMVMTVEKLEIERLKTKYKOLEAKMAQKAE	478				
DB	430	LLKLDKILTEKGGKKAQLLSTPKRERMVMTVEKLEIERLKTKYKOLEAKMAQKAE	489				
QY	479	-EKENH	483				
DB	490	LEKSN	495				
RESULT 8							
QY18KO	ID	Q918KO	PRELIMINARY;	PRT;	651 AA.		
AC	Q918KO;						
DT	01-OCT-2000	(TREMBlrel. 15, Created)					
DT	01-OCT-2000	(TREMBlrel. 15, Last sequence update)					
DT	01-MAR-2004	(TREMBlrel. 26, Last annotation update)					
DE	Chromokinesin Xkid.						
OS	Xenopus laevis (African clawed frog).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;						
OC	Xenopodinae; Xenopus.						
OX	NCBI_TaxID=8355;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	MEDLINE=20419288; PubMed=10966104;						
RX	Funabiki H., Murray A.W.;						
RA	"The Xenopus chromokinesin Xkid is essential for metaphase chromosome						
RT	alignment and must be degraded to allow anaphase chromosome						
RT	movement.";						
RL	Cell 102:411-424(2000).						
DR	EMBL; AF267850; AAF82564.1; --						
DR	HSSP; P33173; I15S.						
DR	GO; GO:0005622; C:intracellular; IEA.						
DR	GO; GO:0005875; C:microtubule associated complex; IEA.						
DR	GO; GO:0005524; F:ATP binding; IEA.						
DR	GO; GO:0003677; F:DNA binding; IEA.						
DR	GO; GO:0003774; F:motor activity; IEA.						
DR	InterPro; IPR010996; DNAPOL_B_N-like.						
DR	InterPro; IPR003583; HHH_1.						
DR	InterPro; IPR001752; kinesin_motor.						
DR	Pfam; PF00225; Kinesin; 1.						
DR	PRINTS; PR00380; KINESINHEAVY.						
DR	SMART; SM00278; HHH1; 2.						
DR	SMART; SM00129; KISC; 1.						
DR	PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.						

Pfam; PF00225; Kinesin; 1.  
DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
DR ATP-binding; Microtubule; Motor protein.  
KW SMART; SM00278; Hhh1; 2.  
SQ SEQUENCE 651 AA; 73091 MW; C4EF79801B603C13 CRC64;

Query Match 58.1%; Score 1437; DB 2; Length 651;  
Best Local Similarity 60.1%; Pred. No. 1.3e-85;  
Matches 292; Conservative 80; Mismatches 100; Indels 14; Gaps 7;  
Mates 292; Conservative 80; Mismatches 100; Indels 14; Gaps 7;

QY 5 RLSKIGATRRPPARVRVAVRLRPFVDGTAGASDP--PCVGRMDSCSLSEIANWRNHQTLLK 63  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 17 RVSMLDQHKKSCARVRVAVRLRPFYMDXDEAKATTVCVGRGLDSQSLEIVNWRNQLETMQ 76  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 64 YQFDAPFYGERSTOODIYAGSVQPIHLRHLEGONASVLAYGPTGAGKTHTMLGSPQPGVI 123  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 77 YQFDAPFYGDSSAQREIYMGVCVCHILPHLLIGONASVFAYGPTGAGKTHTMLGNPNQPGVI 136  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 124 PRALMDLLOLTREGA--EGRPWALSVMYSLEYIQEKVLDLDDPASGDLVIREDCRGNI 181  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 137 PRAVRDQLQNSRTAASAPENENWTYTITMSYVEIYQEKVMDLLEPKNDLPFIRSDKHNI 196  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 182 LIPGLSKQTISSPADPERHFLPASNRRTVGATRLNQRSSRSNAVLLVKVDORERLAPFRQ 241  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 197 LIPGVTKMINSFPADPDEHFIPASQNRTVASTKLNDRSSRSNAVLLIKVKSQQVVPRQ 256  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 242 REGKLYLIDLAGSEDNRRNTGNGRLKESGAINTSFLVKGKVVDALNOGLPRVPYRDSKL 301  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 257 LTGKLYLIDLAGSEDNRRNTGNGRIKESGAINSSLFTLSKVVDALNOGLPRIPYRDSKL 316  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 302 TRLLQDSLGSASHSILLANIAPERFYLDTVSALNFAARSKEVINRNPETNESLOPHALGP 361  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 317 TRLLQDSLGSASHSVMITNIAPEQTYFFDTLTALNFAAKSKOIINKPFSQETTQTVVOPA 376  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 362 VKLSOKE---LLGPPPEAKARGPEEEIGSGPEPMAAPASOKLSPLQKLSMDMPAMLER 418  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 377 MKPREETHIAGSKQKRKSKNDSTE--SSPN-SSMDTAGKQKLN----LATLDPAAVER 429  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 419 LLSDLRLASQSGCAPLLSTPKRERMVLMKTVEEKOLEIERLKTKQKELEAKMLAQKAE 478  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 430 LKLDKILTKEGKKAQALLSTPKRERMALLKKWEESQMEIERLKEQKELEQKAMEAEAR 489  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 479 -EKENH 483  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 490 LEKSN 495

RESULT 9  
Q7ZYL5 PRELIMINARY; PRT; 663 AA.

ID Q7ZYL5 AC Q7ZYL5  
Q7ZYL5; Q7ZYL5;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Similar to kinesin family member 22 (Fragment).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP TISSUE=Embryo;  
RC Klein S., Strausberg R.;  
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
RL EMBL; BC043733; AAA43733.1; -.  
DR HSSP; P33173; I155.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005875; C:microtubule associated complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0003774; F:motor activity; IEA.  
DR InterPro; IPR010996; DNaPol\_B\_like.  
DR InterPro; IPR003583; HHH 1.  
DR InterPro; IPR001752; kinesin motor.

Pfam; PF00225; Kinesin; 1.  
DR PRINTS; PRO0380; KINESINHEAVY.  
DR SMART; SM00129; KISC; 1.  
DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
KW ATP-binding; Microtubule; Motor protein.  
FT NON\_TER 1  
SQ SEQUENCE 663 AA; 74538 MW; 3B98B5F3C12C9C22 CRC64;

Query Match 57.8%; Score 1428; DB 2; Length 663;  
Best Local Similarity 59.4%; Pred. No. 5.1e-85;  
Matches 287; Conservative 82; Mismatches 106; Indels 8; Gaps 4;  
Mates 287; Conservative 82; Mismatches 106; Indels 8; Gaps 4;

QY 5 RLSKIGATRRPPARVRVAVRLRPFVDGTAGASDP--PCVGRMDSCSLSEIANWRNHQTLLK 63  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 30 RVSMLDQHKKSCARVRVAVRLRPFYMDXDEAKATTVCVGRGLDSQSLEIVNWRNQLETMQ 89  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 64 YQFDAPFYGERSTOODIYAGSVQPIHLRHLEGONASVLAYGPTGAGKTHTMLGSPQPGVI 123  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 90 YQFDAPFYGDSSAQREIYMGVCVCHILPHLLIGONASVFAYGPTGAGKTHTMLGNPNQPGVI 149  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 124 PRALMDLLOLTRE--EGAGRPWALSVMYSLEYIQEKVLDLDDPASGDLVIREDCRGNI 181  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 150 PRAVRDQLQNSRTAASAPENENWTYTITMSYVEIYQEKVMDLLEPKNDLPFIRSDKHNI 209  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 182 LIPGLSKQTISSPADPERHFLPASNRRTVGATRLNQRSSRSNAVLLVKVDORERLAPFRQ 241  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 210 LIPGVTKMINSFGDFDEHFIPASQNRTVASTKLNDRSSRSNAVLLIKVKSQQVVPRQ 269  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 242 REGKLYLIDLAGSEDNRRNTGNGRLKESGAINTSFLVKGKVVDALNOGLPRVPYRDSKL 301  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 270 LIGKLYLIDLAGSEDNRRNTGNGRIKESGAINSSLFTLSKVVDALNOGLPRIPYRDSKL 329  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 302 TRLLQDSLGSASHSILLANIAPERFYLDTVSALNFAARSKEVINRNPETNESLOPHALGP 361  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 330 TRLLQDSLGSASHSVMITNIAPEQTYFFDTLTALNFAAKSKOIINKPFSQETTQTVVOPA 389  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 362 VKLSOKE---LLGPPPEAKARGPEEEIGSGPEPMAAPASOKLSPLQKLSMDMPAMLERLLS 421  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 390 MKPREEABATTSRQRKSKTDESTPSNMSSWESTGKRKLN----LASUDSAVVERLLK 445  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 422 LDRLASQSGCAPLLSTPKRERMVLMKTVEEKOLEIERLKTKQKELEAKMLAQKAE-EK 480  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 446 LDKILTKEGKKAQALLSTPKRERMALLKKWEESQMEIERLKEQKELEQKAMEAEARLEK 505  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
QY 481 ENH 483  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 506 SN 508

RESULT 10  
Q919A8 PRELIMINARY; PRT; 650 AA.

ID Q919A8 AC Q919A8  
Q919A8; Q919A8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Kinesin.  
GN Name=kid;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20419289; PubMed=10966105;  
RA Antonio C., Terby I., Wilhelm H., Jones M., Karsenti E., Nebreda A.R.,  
RA Vernos I.;  
RT "Kid, a chromokinesin required for chromosome alignment on the metaphase plate";  
RL Cell 102:425-435(2000).

```

DR EMBL; AJ249841; CAB71799.1; -.
DR HSSP; P31173; 1155.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0008975; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR010996; HNH1_B_N_like.
DR InterPro; IPR003583; HNH1.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR SMART; SM00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00467; KINESIN MOTOR DOMAIN2; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ
SEQUENCE 650 AA; 7307 MW; EA8415BC4B9B0F4A CRC64;
Query Match 57.6%; Score 1423; DB 2; Length 650;
Best Local Similarity 59.4%; Pred. No. 1e-84;
Matches 287; Conservative 81; Mismatches 107; Indels 8; Gaps 4;
QY 5 RLSKIGATRRPPARVRVAVLRPFVDGTAGASDP-PCVGRGMDSCSLEITANRNHNETLK 63
DB 17 RASMLDQHKSSCARVRVAVLRPYMEKEEDKVPACVRLGDSHSLVNNRNQLETHQ 76
QY 64 YQDFAFGERSTODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMLGSPQPGVI 123
DB 77 YQDFAFGDSASQREIYMGVSVCHILPHLLIGQNASVFAYGPTGAGKTHMLGNPDQPGVI 136
QY 124 PRALMDLQLTR--EGAGRPWALSVTMSYLIYQEKVLDLDPASGLDVIREDCRGNI 181
DB 137 PRAVRELLQMTMAASAPENENNTHTITSYVEIYQEKVMDLLEPKNKDLPREDKDHNI 196
QY 182 LIPLGSKPISFADFEHFLPASRNTVGATRLNORSSSHAVLVKVDQERLAPFRQ 241
DB 197 LIPGVTLTINSFGDFDAHFIPASQNTVASTKLNDRSSSHAVLVKVKQSQVAFPRQ 256
QY 242 REGKYLIDLAGSDNRNRTGKGLRKESGAINSTSLFVLGVVDALNQGLPRVPYRDSKL 301
DB 257 LIGKYLIDLAGSDNRNRTGKGLRKESGAINSTSLFVLGVVDALNQGLPRVPYRDSKL 316
QY 302 TRLLQDSLGSAHSLTANTAPERFVLTVAINFARSKVINPPTNESIQHALGP 361
DB 317 TRLLQDSLGSAHSLTANTAPERFVLTVAINFARSKVINPPTNESIQHALGP 376
QY 362 VLKSKELLGPPKAKRGPEEIEIGSPPEMAAPASASOKLSPLQLSSMDPAMLERLLS 421
DB 377 MKRPREAEATTSRQRKSKTDSSTSPNSMESTGKRLN---LASLDSAVVERLLK 432
QY 422 LDRLASQGSQCAPLLSTPKRVMVLTVEEKDLEIRLTKQKELEAKMLQAKAB-EK 480
DB 433 LDKILTEKGKEAQLLSTPKRVMVLTVEEKDLEIRLTKQKELEAKMLQAKAB-EK 492
QY 481 ENH 483
DB 493 SNN 495
RESULT 11
QSGPGO PRELIMINARY; PRT; 631 AA.
AC Q6GPGO;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;

```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RC SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RC SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC073177; AAH73177.1; -.
DR InterPro; IPR001752; kinesin motor.
DR InterPro; IPR010994; Ruva_2_like.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Hypothetical protein; Microtubule; Motor protein.
SQ
SEQUENCE 631 AA; 71084 MW; 1158C9AB86EBD3F3 CRC64;
Query Match 57.5%; Score 1422; DB 2; Length 631;
Best Local Similarity 60.6%; Pred. No. 1.2e-84;
Matches 285; Conservative 78; Mismatches 99; Indels 8; Gaps 4;
QY 18 ARVRVAVRLRPFDVGTAGASDP-PCVGRGMDSCSLEITANRNHNETLKQYDFAGYGERSTQ 76
DB 11 ARVRVAVRLRPYMEKEEDKVPACVRLGDSHSLVNNRNQLETHQYDFAGYGERSTQ 70
QY 77 QDIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMLGSPQPGVIPALMDLQLTR- 135
DB 71 REIYMGVSVCHILPHLLIGQNASVFAYGPTGAGKTHMLGNPDQPGVIPRAVRELLQMTRM 130
QY 136 -EGAGRPWALSVTMSYLIYQEKVLDLDPASGLDVIREDCRGNILPGLSQKPISSF 194
DB 131 AASAPENENNTHTITSYVEIYQEKVMDLLEPKNKDLPREDKDHNLIPGVTLKTNF 190
QY 195 ADPERHFLPASRNTVGATRLNORSSSHAVLVKVDQERLAPFRQREGKLYLIDLAGS 254
DB 191 GDFDEHFIPASQNTVASTKLNDRSSSHAVLVKVKQSQVAFPRQLKLYLIDLAGS 250
QY 255 EDNRRTGNKGLRKESGAINSTSLFVLGVVDALNQGLPRVPYRDSKLTRLODSLGSAH 314
DB 251 EDNRRTGNQGIKRLKESGAINSTSLFVLGVVDALNQGLPRVPYRDSKLTRLODSLGSAH 310

```

```

Qy 315 SILIANTIPERRFYLDVTSALNFAARKEVINRFTNESLQPHALGPVKLSQKELGPPE 374
Db 311 SWMITIANPEQTYFDLTALNFAAKSKQIINKPFSRETTQVAPAKMRPREAEATTS 370
Qy 375 AKRARGPEERIGSPPEWAPASAKLSPLQKLSMDPAMLERLLSLDRLLASQSGQA 434
Db 371 SRQKRSKTDSTESPNSSMESTGKRLN----LASLDSAVVERLLKDLTKLTTEKGRKEA 426
Qy 435 PLLSTPKRRMLVMTVEEKLEIERLKTQKLEAKMLQAKAE-EXENH 483
Db 427 QLLSTPKRRMALLKKWEESQMEIERLKEKQKLEQKAMEARLEKSN 476

RESULT 12
O60845 PRELIMINARY; PRT; 198 AA.
AC O60845;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to kinesin-like DNA binding protein (KID).
GN Name=A-328A3.2;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=94425270; PubMed=10493829;
RA Loftus B.-J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.B., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
DR EMBL; AC002301; AAC08709.1; -.
DR HSP; P33173; I161.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 198 AA; 21567 MW; 1A7FDA95B5A03B37 CRC64;

Query Match 33.8%; Score 836; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.7e-47;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCVRGNDSCSLIANWRNHQET 61
Db 26 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCVRGNDSCSLIANWRNHQET 85
Qy 62 LKQFDIFYGERSTQDIYAGSVQPIRLHLEGGNASVLAGPTGACKTHMLGSPQPG 121
Db 86 LKQFDIFYGERSTQDIYAGSVQPIRLHLEGGNASVLAGPTGACKTHMLGSPQPG 145

Qy 122 VIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKV 160
Db 146 VIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKV 184

RESULT 13
O35232 PRELIMINARY; PRT; 148 AA.
AC O35232;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kinesin motor protein KIF22 (Fragment).

```

```

GN Name=Kif22;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=97480723; PubMed=9339368;
RA Yang Z., Hanlon D.W., Marszalek J.R., Goldstein L.S.;
RT "Identification, partial characterization, and genetic mapping of
RT kinesin-like protein genes in mouse.";
RL Genomics 45:123-131(1997).
DR EMBL; AF013119; AAC39968.1; -.
DR MGD; MGI:109233; Kif22.
DR GO; GO:0000785; C:chromatin; IDA.
DR GO; GO:0005819; C:spindle; IDA.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW ATP-binding; Microtubule; Motor protein.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 16596 MW; 0950C265C3B63EB7 CRC64;

Query Match 28.5%; Score 704; DB 2; Length 148;
Best Local Similarity 92.6%; Pred. No. 1.7e-38;
Matches 137; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 108 GKTHMLGSPQGVIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKVLDLDDPA 167
Db 1 GKTHMLGSPQGVIPRALMDLLQLTREGAAGRPWALSVTMSYLEIYQEKVLDLDDPA 60
Qy 168 SGDLVIREDCRGNILIPGLSQPISSPADFERHPLPASRNTVGATRLNQSSSHAVLL 227
Db 61 SGDLVIREDCRGNILIPGLTQKPIPSDFEQHPLPASRNTVGATRLNQSSSHAVLL 120
Qy 228 VKVDQERLERLAPFRQREGKYLIDLAGSE 255
Db 121 VKVDQERLERLAPFRQREGKYLIDLAGSE 148

RESULT 14
Q6L512 PRELIMINARY; PRT; 584 AA.
AC Q6L512;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative kinesin-like DNA binding protein.
GN Name=OJ1281_H05.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.-C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117265; AAT39162.1; -.
DR InterPro; IPR010996; DNAPol_B_N_like.
DR InterPro; IPR000445; HhH.
DR InterPro; IPR001752; Kinesin_motor.
DR Pfam; PF00633; HHH; 1.
DR Pfam; PF00225; Kinesin; 1.

```

[illegible]

Search completed: November 5, 2004, 18:38:43  
Job time : 73.979 secs

```

DR PRINTS; PRO0380; KINESINHEAVY.
DDR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR KW ATP-binding; Microtubule; Motor protein.
DR SQ SEQUENCE 584 AA; 64124 MW; EF829C3A15P9190F CRC64;

Query Match 25.6%; Score 633; DB 2; Length 584;
Best Local Similarity 36.1%; Pred.No. 5.1e-33;
Matches 166; Conservative 155; Indels 50; Gaps 15;

QY 6 LSKIGATRRPPARVAVRLRPVDCGTAGASDPCCVRGMDS--CSLEIANWENHOETLK- 63
DB 1 MATAAATQSQP---VRVVLVRPHLPSEANSEAPCVLLGSHPGGVEVTVQLKDQVTSRN 57
QY 64 --YQPDAPYGRSTQDDIYAGSVQPIRLHLEQONASVLAGYGTGACKTHTWLMSPEQPG 121
DB 58 ECKYKLDAPFGQESRVCEIFDQVESAIVPGIFEGFNATVFAYGATGSGKTYTMOGTEDLFG 117
QY 122 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIQYOEKVLDDLDPDASGLDVIREDCRGNI 181
DB 118 LIPLAVSTVLAICT-----GTWC-SVEISVYEVYMERCYDLELPEKAREIMVLDDKGNL 170
QY 182 LIPGLSQKPIGSFADFERHFPLPASRNTVGNATRLNQSRSHAVLLVKVDQRERLAPFQ 241
DB 171 QLKGLAVPVRSLBEFHEIYISIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 223
QY 242 REGKLYLIDLAGSEDRNTGNKGLRKESGAINSTSLFVLGVVDALNOGLPRVPYRDSKL 301
DB 224 VVGKLNLIIDLAGNDRRTCTNEGIRLQESAKINOSLFALSNVISALNKKPEAIPYRESKL 283
QY 302 TRLLQDSLGGSAAHSILIANIAPERRFRYDVTYSALNFAARSKEVINRPFNTNESLPHALGP 361
DB 284 TRILQDSLGNSHAVMIACLNPVE--YQEAHVHTVSLAARSHRVTNH--MSSASKQETPKDK 340
QY 362 VKLSQK-----ELLGPPEA-KRAGCPBEEIGSPPEWAAP-----ASASOKLSPLQKL 408
DB 341 VDMEAKRLAWLESKGTSTIORMDG-----LLSPNAIKTFLMSHKKQASGRVSGEKA 395
QY 409 SSMDFPAMLERLL--SLDRLLAS-----QGSQAPILLSTPKR 442
DB 396 MNQDGGKIKKVI-FDASAARTPAENPHREGTO--DIWNTTKK 433

```

Query Match 25.6%; Score 633; DB 2; Length 584;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 20.0023 Seconds  
(without alignments)  
1697.543 Million cell updates/sec

Title: US-10-797-893-4  
Perfect score: 2589  
Sequence: 1 MPAAGGSGTQRRREMAAAS.....LEAKMLAKAEKENHCPM 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/6C\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	100.0	512	3	US-09-724-224-4
2	2589	100.0	512	4	US-10-093-317-4
3	2577	99.5	665	4	US-09-595-684B-35
4	2467	95.3	487	3	US-09-724-224-8
5	2467	95.3	487	4	US-10-093-317-8
6	1873	72.3	370	3	US-09-724-224-2
7	1873	72.3	370	4	US-10-093-317-2
8	1763	68.1	346	3	US-09-724-224-6
9	1763	68.1	346	4	US-10-093-317-6
10	573.5	22.2	1234	4	US-09-592-054-8
11	568.5	22.0	473	4	US-09-592-054-6
12	568.5	22.0	522	4	US-09-592-054-4
13	561.5	21.7	1232	4	US-09-592-054-2
14	561	21.7	1388	3	US-09-592-191-2
15	561	21.7	1388	3	US-09-723-262-2
16	561	21.7	1388	3	US-09-723-219-2
17	565.5	21.5	1066	3	US-09-541-782-8
18	556.5	21.5	1066	4	US-09-723-820-8
19	556.5	21.5	1066	4	US-10-270-085-8
20	554	21.4	1279	3	US-09-724-517-2
21	554	21.4	1279	4	US-09-641-807A-2
22	554	21.4	1279	4	US-09-723-096-2
23	553.5	21.4	1231	4	US-09-595-684B-23
24	551	21.3	2954	4	US-09-150-867-1
25	550	21.2	864	4	US-09-883-096-2
26	549	21.2	355	3	US-09-724-511-4
27	549	21.2	355	4	US-09-723-097-4

28	549	21.2	355	4	US-09-632-344-4	Sequence 4, Appli
29	549	21.2	367	3	US-09-724-511-2	Sequence 2, Appli
30	549	21.2	367	4	US-09-723-097-2	Sequence 2, Appli
31	549	21.2	367	4	US-09-632-344-2	Sequence 2, Appli
32	546.5	21.1	1103	3	US-09-162-373-1	Sequence 1, Appli
33	546.5	21.1	1103	3	US-09-467-946-1	Sequence 25, Appli
34	544.5	21.0	935	4	US-09-914-259-25	Sequence 6, Appli
35	543	21.0	409	3	US-09-572-191-6	Sequence 6, Appli
36	543	21.0	409	3	US-09-723-262-6	Sequence 6, Appli
37	543	21.0	409	3	US-09-723-219-6	Sequence 23, Appli
38	543	21.0	928	4	US-09-914-259-23	Sequence 39, Appli
39	541	20.9	1690	4	US-09-595-684B-39	Sequence 6, Appli
40	535	20.7	513	4	US-09-724-519-6	Sequence 6, Appli
41	535	20.7	513	4	US-09-592-037-6	Sequence 6, Appli
42	535	20.7	513	4	US-09-438-156B-6	Sequence 2, Appli
43	533.5	20.6	1057	4	US-09-428-156B-2	Sequence 29, Appli
44	532	20.5	1056	4	US-09-595-684B-29	Sequence 10, Appli
45	532	20.5	1057	3	US-09-541-782-10	

## ALIGNMENTS

RESULT 1  
US-09-724-224-4  
; Sequence 4, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724, 224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597, 292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-224-4

Query Match	100.0%;	Score 2589;	DB 3;	Length 512;
Best Local Similarity	100.0%;	Pred. No. 1.8e-233;		
Matches 512;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MPAAGGSGTQRRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPPFVDGTAG	60
Db	1	MPAAGGSGTQRRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPPFVDGTAG	60
Qy	61	ASDPPCVGRMGDSCLSEIANWRNHQETLKYQF	DAFYGERSTQDDIYAGSVQFILRHLEGG	120
Db	61	ASDPPCVGRMGDSCLSEIANWRNHQETLKYQF	DAFYGERSTQDDIYAGSVQFILRHLEGG	120
Qy	121	NASVLAYGPTGAGKTHMLGSPQPGV	IPRALMDLLQITREEGAGRPWALSVMYSYLEI	180
Db	121	NASVLAYGPTGAGKTHMLGSPQPGV	IPRALMDLLQITREEGAGRPWALSVMYSYLEI	180
Qy	181	YQEKVLDLLDPASGDLVIREDCRGNILIPGL	SQKPISSFADFERRHFLPASNNRTVGATRL	240
Db	181	YQEKVLDLLDPASGDLVIREDCRGNILIPGL	SQKPISSFADFERRHFLPASNNRTVGATRL	240
Qy	241	NQSRSSSHAVLLVKVDQERLAPPRQREGKLYL	IDLAGSEDNRRRTGNKGLRKESGAINT	300
Db	241	NQSRSSSHAVLLVKVDQERLAPPRQREGKLYL	IDLAGSEDNRRRTGNKGLRKESGAINT	300
Qy	301	SLFVLGKVDALNOGLPRVYRDSKLTLLQDS	LGSGSAHSILIANIAPERFYLDTVSAL	360
Db	301	SLFVLGKVDALNOGLPRVYRDSKLTLLQDS	LGSGSAHSILIANIAPERFYLDTVSAL	360
Qy	361	NFAARSEVINRPTNESLQPHALGPVKLSQKEL	LGPPPEAKRGARPEEEIGSPPEPMAAP	420

Db 361 NFAARKEVINRFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEWAP 420  
QY 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKD 480  
Db 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKD 480  
QY 481 LEIERLTKTKQKELEAKMLAQKAEKENHCPTM 512  
Db 481 LEIERLTKTKQKELEAKMLAQKAEKENHCPTM 512

RESULT 2  
US-10-093-317-4  
; Sequence 4, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; CURRENT FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Human  
US-10-093-317-4

Query Match  
Best Local Similarity 100.0%; Score 2589; DB 4; Length 512;  
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPAAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVVLRPFVDGTAG 60  
Db 1 MPAAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVVLRPFVDGTAG 60  
QY 61 ASDPPCVRGMDSCSLEIANRNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGG 120  
Db 61 ASDPPCVRGMDSCSLEIANRNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGG 120  
QY 121 NASVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEI 180  
Db 121 NASVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEI 180  
QY 181 YQSKVLDLDDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNRVTGATRL 240  
Db 181 YQSKVLDLDDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNRVTGATRL 240  
QY 241 NQSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDNRRTGNKGLRKESGAIN 300  
Db 241 NQSSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDNRRTGNKGLRKESGAIN 300  
QY 301 SLFVLGVVDALNOGLPRVPYRDSKLTRELLOSLGSGSAHSILIANIAPERFYLDTVSAL 360  
Db 301 SLFVLGVVDALNOGLPRVPYRDSKLTRELLOSLGSGSAHSILIANIAPERFYLDTVSAL 360  
QY 361 NFAARKEVINRFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEWAP 420  
Db 361 NFAARKEVINRFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEWAP 420  
QY 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKD 480  
Db 421 ASASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKD 480  
QY 481 LEIERLTKTKQKELEAKMLAQKAEKENHCPTM 512  
Db 481 LEIERLTKTKQKELEAKMLAQKAEKENHCPTM 512

RESULT 3  
US-09-595-684B-35  
; Sequence 35, Application US/09595684B  
; Patent No. 6544766  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Ohashi, Cara  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Vaisberg, Eugeni  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Yu, Ming  
; TITLE OF INVENTION: Human kinesins and methods of producing  
; TITLE OF INVENTION: and purifying human kinesins  
; FILE REFERENCE: cytop036  
; CURRENT APPLICATION NUMBER: US/09/595,684B  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Human  
US-09-595-684B-35

Query Match  
Best Local Similarity 100.0%; Score 2577; DB 4; Length 665;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVVLRPFVDGTAGS 62  
Db 2 AAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPARVAVVLRPFVDGTAGS 61  
QY 63 DPPCVRGMDSCSLEIANRNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGGNA 122  
Db 62 DPPCVRGMDSCSLEIANRNHQETLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGGNA 121  
QY 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182  
Db 122 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 181  
QY 183 EKVLDDLDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNRVTGATRLNQ 242  
Db 182 EKVLDDLDPASGDLVIREDCRGNILIPGLSOKPISSFADPERHFLPASRNRVTGATRLNQ 241  
QY 243 RSRSSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDNRRTGNKGLRKESGAIN 302  
Db 242 RSRSSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDNRRTGNKGLRKESGAIN 301  
QY 303 FVLGVVDALNOGLPRVPYRDSKLTRELLOSLGSGSAHSILIANIAPERFYLDTVSALNF 362  
Db 302 FVLGVVDALNOGLPRVPYRDSKLTRELLOSLGSGSAHSILIANIAPERFYLDTVSALNF 361  
QY 363 AARSKEVINRFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAPAS 422  
Db 362 AARSKEVINRFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAPAS 421  
QY 423 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDLE 482  
Db 422 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDLE 481  
QY 483 IERLTKTKQKELEAKMLAQKAEKENHCPTM 512  
Db 482 IERLTKTKQKELEAKMLAQKAEKENHCPTM 511

RESULT 4  
US-09-724-224-8  
; Sequence 8, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/09/724,224  
 ; CURRENT FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 487  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-724-224-8

Query Match 95.3%; Score 2467; DB 3; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-222;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 27 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 86  
 Db 2 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 61  
 QY 87 LKYOQDFPYGERSSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQPG 146  
 Db 62 LKYOQDFPYGERSSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQPG 121  
 QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206  
 Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181  
 QY 207 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 266  
 Db 182 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 241  
 QY 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 326  
 Db 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 301  
 QY 327 TRLLQDSLGSSAHSILIANIAPERFFYLDTVSALNFAARSKEVINRPFNLSLOPHALGP 386  
 Db 302 TRLLQDSLGSSAHSILIANIAPERFFYLDTVSALNFAARSKEVINRPFNLSLOPHALGP 361  
 QY 387 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 446  
 Db 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
 QY 447 LDRLLASQGSQCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKQLEAKMLAQKAEKE 506  
 Db 422 LDRLLASQGSQCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKQLEAKMLAQKAEKE 481  
 QY 507 NHCPTM 512  
 Db 482 NHCPTM 487

RESULT 5  
 US-10-093-317-8  
 ; Sequence 8, Application US/10093317  
 ; Patent No. 6762043  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/093,317  
 ; CURRENT FILING DATE: 2002-03-06  
 ; PRIOR APPLICATION NUMBER: 09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8

; LENGTH: 487  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-10-093-317-8  
 Query Match 95.3%; Score 2467; DB 4; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-222;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 27 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 86  
 Db 2 GRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET 61  
 QY 87 LKYOQDFPYGERSSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQPG 146  
 Db 62 LKYOQDFPYGERSSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQPG 121  
 QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206  
 Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181  
 QY 207 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 266  
 Db 182 LIPGLSOKPISSPADFERHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPPRQ 241  
 QY 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 326  
 Db 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGLPRVPYRDSKL 301  
 QY 327 TRLLQDSLGSSAHSILIANIAPERFFYLDTVSALNFAARSKEVINRPFNLSLOPHALGP 386  
 Db 302 TRLLQDSLGSSAHSILIANIAPERFFYLDTVSALNFAARSKEVINRPFNLSLOPHALGP 361  
 QY 387 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 446  
 Db 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
 QY 447 LDRLLASQGSQCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKQLEAKMLAQKAEKE 506  
 Db 422 LDRLLASQGSQCAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKQLEAKMLAQKAEKE 481  
 QY 507 NHCPTM 512  
 Db 482 NHCPTM 487

RESULT 6  
 US-09-724-224-2  
 ; Sequence 2, Application US/09724224  
 ; Patent No. 6387644  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/09/724,224  
 ; CURRENT FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 370  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-724-224-2

Query Match 72.3%; Score 1873; DB 3; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-166;  
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 AAGGSTQQRREMAAASAAISAGRCRLSKIGATRRPPPPARVRVAVRLRPFVDGTAGAS 62

Db 2 AAGGTOORREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 61  
QY 63 DPCVCGMDSCLSIANWRNHQETLKYQDFAGYGERSTQODIYAGSVQPIRLHLLGQNA 122  
Db 62 DPCVCGMDSCLSIANWRNHQETLKYQDFAGYGERSTQODIYAGSVQPIRLHLLGQNA 121  
QY 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIQ 182  
Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIQ 181  
QY 183 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFPERHFLPASRNRVTGATRLNQ 242  
Db 182 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFPERHFLPASRNRVTGATRLNQ 241  
QY 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLKESGAINSTSL 302  
Db 242 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLKESGAINSTSL 301  
QY 303 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 362  
Db 302 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 361  
QY 363 AARSKEVIN 371  
Db 362 AARSKEVIN 370  
RESULT 7  
US-10-093-317-2  
; Sequence 2, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR FILING DATE: 09/724,224  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Human  
US-10-093-317-2  
Query Match 72.3%; Score 1873; DB 4; Length 370;  
Best Local Similarity 100.0%; Pred. No. 1.1e-166; Indels 0; Gaps 0;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 AAGGTOORREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 62  
Db 2 AAGGTOORREMAAASAAISGAGCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGAS 61  
QY 63 DPCVCGMDSCLSIANWRNHQETLKYQDFAGYGERSTQODIYAGSVQPIRLHLLGQNA 122  
Db 62 DPCVCGMDSCLSIANWRNHQETLKYQDFAGYGERSTQODIYAGSVQPIRLHLLGQNA 121  
QY 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIQ 182  
Db 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRPPWALSVTMSYLEIQ 181  
QY 183 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFPERHFLPASRNRVTGATRLNQ 242  
Db 182 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFPERHFLPASRNRVTGATRLNQ 241  
QY 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLKESGAINSTSL 302  
Db 242 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLKESGAINSTSL 301  
QY 303 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 362

Db 302 FVLGKVDALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 361  
QY 363 AARSKEVIN 371  
Db 362 AARSKEVIN 370  
RESULT 8  
US-09-724-224-6  
; Sequence 6, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR FILING DATE: 09/597,292  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-224-6  
Query Match 68.1%; Score 1763; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 2e-156; Indels 0; Gaps 0;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 27 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86  
Db 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
QY 87 LKQDFAGYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPQPG 146  
Db 62 LKQDFAGYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPQPG 121  
QY 147 VIPRALMDLLQLTREBAGRPPWALSVTMSYLEIQEKVLDLDDPASGDLVIREDCRGN 206  
Db 122 VIPRALMDLLQLTREBAGRPPWALSVTMSYLEIQEKVLDLDDPASGDLVIREDCRGN 181  
QY 207 LIPGLSQPISSFADFPERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 266  
Db 182 LIPGLSQPISSFADFPERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 241  
QY 267 REGKLYLIDLAGSEDNRRTGNKGLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 326  
Db 242 REGKLYLIDLAGSEDNRRTGNKGLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 301  
QY 327 TRLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371  
Db 302 TRLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
RESULT 9  
US-10-093-317-6  
; Sequence 6, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR FILING DATE: 09/724,224  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-6

Query Match      68.1%; Score 1763; DB 4; Length 346;
Best Local Similarity 100.0%; Pred. No. 2e-156;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCRLSKIGATRRPPPARVRVAVRLRPFDVTAGASDPPCVRGMDSCSLEIANWRNHQET 86
Db 2 GRCRLSKIGATRRPPPARVRVAVRLRPFDVTAGASDPPCVRGMDSCSLEIANWRNHQET 61

QY 87 LKYQFDAPYGRSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPROPG 146
Db 62 LKYQFDAPYGRSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPROPG 121

QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 206
Db 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181

QY 207 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRHAVLLVKVDORERLAPRQ 266
Db 182 LIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRHAVLLVKVDORERLAPRQ 241

QY 267 REGKYLIDLAGSEDNRRTGNKGLRKESGAINSTSLFVLGVKVDALNQGRLPRVYRDSKL 326
Db 242 REGKYLIDLAGSEDNRRTGNKGLRKESGAINSTSLFVLGVKVDALNQGRLPRVYRDSKL 301

QY 327 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371
Db 302 TRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

RESULT 10
US-09-592-054-8
; Sequence 8, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592.054
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1234
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-8

Query Match      22.2%; Score 573.5; DB 4; Length 1234;
Best Local Similarity 32.2%; Pred. No. 3e-44;
Matches 156; Conservative 97; Mismatches 173; Indels 59; Gaps 13;

QY 45 VRVAVRLRPFDVTAGASDPPC-----VRGMDSCSLEIANWRNHQETLKYOPDAP 94
Db 10 VRVALRCPLVPKEISEGCMQCLSFVPGTQVVVGTDK-----SFTYDFV 54

QY 95 YGERSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPROPG--GV 147
Db 55 FDPCTEQUEEVFNKAVAPLIKGIFKYNATVLAIVGQTSKTYSMGGAYTAQEENPTVI 114

QY 148 IPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205
Db 115 IPR-----VIQLFKEDIKSDPFTLKVSYLYNEEILDLCPSREKAQINIREDPKEG 170

206 ILIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRHAVLLVKVDORERLAPR 265
171 IKIVGLEKTVLVALDVTVCLEQNNSTVASTAMNQSRRSHAFTISLEQCKSKDKNS 230
266 QREGKYLIDLAGSEDNRRTGNKGLRKESGAINSTSLFVLGVKVDALNQGRLPR--VPYRD 323
231 SFRSKLHLVDLAGSERQKTKAEGDRLKEGINIRGLCLGNVISALGDDKKGSFVPYRD 290
324 SKLTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLQPHA 383
291 SKLTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNESLQPHA 348
384 LGPVKLSQKELLGPPPEAKRARPPEEIEIGSPPEMAAPASQKLSPLQKLSMDPAMLER 443
349 AELNHLAQ-----QVQQLQVLLLOAHGGTLPQSIINAEPSN---LQSLMEKQNSLVEE 398
444 LLSLDRLLAGSGSGAPLLSTPKERMVLMKTVEEK-DLEIERLKTQK-ELEAKMLAQK 501
399 NEKLSRCLSKRAAGTQAML-----ERILTEQVNEKLNKLEELRQHAACKLDLQKLVET 453
502 AEEKE 506
454 LEDQE 458

RESULT 11
US-09-592-054-6
; Sequence 6, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592.054
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-6

Query Match      22.0%; Score 568.5; DB 4; Length 473;
Best Local Similarity 32.3%; Pred. No. 1.9e-44;
Matches 159; Conservative 97; Mismatches 162; Indels 75; Gaps 15;

QY 45 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLKY 89
Db 8 VRVALRCPLVPKEISEGCMQCLSFVPG-----EPQVVVGTDK-----SF 47

QY 90 QDFAFYGERSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPROPG 144
Db 48 TYDFVDPDSTEQBEVENTAVAPLIKGIFKYNATVLAIVGQTSKTYSMGGAYTAQEEN 107

QY 145 P--GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIRE 200
Db 108 PTGVGIVP-----VIQLFKEDIKSDPFTLKVSYLYNEEILDLCPSREKAQINIRE 163

QY 201 DCRGNILIPGLSQPISSPADFERHFLPASNRRTVGATRLNORSSRHAVLLVKVDORER 260
Db 164 DPKEGKIVGLEKTVLVALDVTVCLEQNNSTVASTAMNQSRRSHAFTISLEQCKK 223

QY 261 LAFPRQREGKYLIDLAGSEDNRRTGNKGLRKESGAINSTSLFVLGVKVDALNQGRLPR-- 318
Db 224 SDKNSFRSKLHLVDLAGSERQKTKAEGDRLKEGINIRGLCLGNVISALGDDKKGGF 283
319 VPYRDSKLTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNES 378
```

Db 284 VPYRDKLRLQLDLSGNSHTLMIACVSPADSNLEETLNTLYADRAKKNKPIVND 343  
QY 379 LQPHALGPVK--LSQKELLGPPPAKARGPEEEIGSPPEMAAPASAKLSP-LQKLS 435  
Db 344 PQTAEHLHLKQVQQLQVL-----LLOAHGG-----TLPGSITVPESENLSLME 388  
QY 436 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKTKQK-EL 493  
Db 389 KNQSLVEENEKLSRGLSEAGQTAQML-----ERIIWTEQANEKNMAKLEELRQHAACKL 443  
QY 494 EAKMLAQKAEKE 506  
Db 444 DLQKLVTLEDQE 456

RESULT 12  
US-09-592-054-4  
; Sequence 4, Application US/09592054  
; Patent No. 6440684  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Finer, Jeffrey  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Wood, Kenneth  
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1016  
; CURRENT APPLICATION NUMBER: US/09/592,054  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 522  
; TYPE: PRT  
; ORGANISM: Human  
US-09-592-054-4

Query Match 22.0%; Score 568.5; DB 4; Length 522;  
Best Local Similarity 32.3%; Pred. No. 2.3e-44;  
Matches 159; Conservative 97; Mismatches 162; Indels 75; Gaps 15;

QY 45 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLKY 89  
Db 30 VRVALRCRPLVPKEISEGCMCLSFVG-----EPQVVVGTDK-----SF 69  
QY 90 QPDAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHTMLGS-----PQ 144  
Db 70 TYDFVDPSTEQEEVENTAVAPLIKGVFKGYNATVLAYGQTSGKTYSMGGAYTAQENE 129  
QY 145 P--GVIPRALMDLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 200  
Db 130 PTVGVIPLR-----VIQLLFKEIDKSDFFETLKVSYLEYINEEILDLCCPSREKAQINIRE 185  
QY 201 DCRGNLILPGLSQKPISSFADFERHFLPASRNRVTGATRLNORSSRSHAVLLVKVDQER 260  
Db 186 DPKEGKIVGLTEKTVLVALDVTVCLEQGNSTVASTAMNSQSSRSHAILTISLEQK 245  
QY 261 LAPFRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTLFLVKVVDALNQGLPR-- 318  
Db 246 SDKNSSFRSKLHLVDLAGSERQKTKAEGDLRKEGININRGLCLGNVISALGDDKKGGF 305  
QY 319 VPYRDKLRLQLDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNES 378  
Db 306 VPYRDKLRLQLDLSGNSHTLMIACVSPADSNLEETLNTLYADRAKKNKPIVND 365  
QY 379 LQPHALGPVK--LSQKELLGPPPAKARGPEEEIGSPPEMAAPASAKLSP-LQKLS 435  
Db 366 PQTAEHLHLKQVQQLQVL-----LLOAHGG-----TLPGSITVPESENLSLME 410  
QY 436 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKTKQK-EL 493  
Db 411 KNQSLVEENEKLSRGLSEAGQTAQML-----ERIIWTEQANEKNMAKLEELRQHAACKL 465

QY 494 EAKMLAQKAEKE 506  
Db 466 DLQKLVTLEDQE 478

RESULT 13  
US-09-592-054-2  
; Sequence 2, Application US/09592054  
; Patent No. 6440684  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Finer, Jeffrey  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Wood, Kenneth  
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1016  
; CURRENT APPLICATION NUMBER: US/09/592,054  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1232  
; TYPE: PRT  
; ORGANISM: Human  
US-09-592-054-2

Query Match 21.7%; Score 561.5; DB 4; Length 1232;  
Best Local Similarity 32.0%; Pred. No. 4e-43;  
Matches 158; Conservative 97; Mismatches 163; Indels 75; Gaps 15;

QY 45 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLKY 89  
Db 10 VRVALRCRPLVPKEISEGCMCLSFVG-----EPQVVVGTDK-----SF 49  
QY 90 QPDAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHTMLGS-----PQ 144  
Db 50 TYDFVDPSTEQEEVENTAVAPLIKGVFKGYNATVLAYGQTSGKTYSMGGAYTAQENE 109  
QY 145 P--GVIPRALMDLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 200  
Db 110 PTVGVIPLR-----VIQLLFKEIDKSDFFETLKVSYLEYINEEILDLCCPSREKAQINIRE 165  
QY 201 DCRGNLILPGLSQKPISSFADFERHFLPASRNRVTGATRLNORSSRSHAVLLVKVDQER 260  
Db 166 DPKEGKIVGLTEKTVLVALDVTVCLEQGNSTVASTAMNSQSSRSHAILTISLEQK 225  
QY 261 LAPFRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTLFLVKVVDALNQGLPR-- 318  
Db 226 SDKNSSFRSKLHLVDLAGSERQKTKAEGDLRKEGININRGLCLGNVISALGDDKKGGF 285  
QY 319 VPYRDKLRLQLDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNES 378  
Db 286 APYRDKLRLQLDLSGNSHTLMIACVSPADSNLEETLNTLYADRAKKNKPIVND 345  
QY 379 LQPHALGPVK--LSQKELLGPPPAKARGPEEEIGSPPEMAAPASAKLSP-LQKLS 435  
Db 346 PQTAEHLHLKQVQQLQVL-----LLOAHGG-----TLPGSITVPESENLSLME 390  
QY 436 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKTKQK-EL 493  
Db 391 KNQSLVEENEKLSRGLSEAGQTAQML-----ERIIWTEQANEKNMAKLEELRQHAACKL 445  
QY 494 EAKMLAQKAEKE 506  
Db 446 DLQKLVTLEDQE 458

RESULT 14  
US-09-572-191-2  
; Sequence 2, Application US/09572191  
; Patent No. 6355466  
; GENERAL INFORMATION:

```

; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
; US-09-572-191-2

```

```

Query Match      21.7%; Score 561; DB 3; Length 1388;
Best Local Similarity 30.9%; Pred. No. 5.4e-43;
Matches 162; Conservative 97; Mismatches 192; Indels 74; Gaps 15;

Qy      45 VRVAVRLRPFDV--GTAGASDPPPCVRGMDSCSLEIANRNHQTETKQDFAFYGERSTQ 102
Db      27 IKVVRIRPPAERSGSADGEONLCLSVLSSTSLR---HSPNPKTTFDHDVADVDITQE 83

Qy      103 DIYAGSVQPIRLHLELEGONASVLAYGPTGAGKTHMLGSPREP-----GVIPRA---L 152
Db      84 SVFATVAKSIVSCSMGVNGTIFAYGQTGSKTFTWNGPSESDNFSNLRGVIPRSFEYL 143

Qy      153 MDLLQLTREGEAGRPWALSVTMGLYIYQKVLDDLPASGDLVIREDCRGNLIPGLS 212
Db      144 FSLIDREKEKAGAGKSPFLCK--CSFIEIYNEQIYDLDLSASAGLYLREHIKKGVFVVGAV 201

Qy      213 QKPISSEADFERHFLPASRNRTVGATRLNQRSSSHAVLLVKVDORERLAPP-RQREGKL 271
Db      202 EQVVTSAEAYQVLSGCGWRNRVASTSMNRESSSHAVFTITTESMEKSEINVRTSL 261

Qy      272 YLIDLAGSENRRTGNKGLRKESGAINTSIFVLGKVVDAL---NOGLPRVPYRDSKLT 327
Db      262 NLVDLAGSERQKQTHAEGMRLEKAGNINRSLSCLGVITLVDVGNKGQRHVCYRDSKLT 321

Qy      328 RLQDSLGGSAHSLIANIAPERFYLDTVSALNFAARSKEVINRPTNESLOPHALGPV 387
Db      322 FLRDSLGNAKTAIIANVHPGRCFCGETLSTLNFAQRAKLIKNKAVVNEDTQ-----GNV 377

Qy      388 KLSQ-----KELIG-----PEAKRARGPREE---EIGSPPEMAAPASASOKLSPLQ 431
Db      378 SQLQAEVKRLKEQLAELASGQTPEPESFLTRDKKNTMYEQEAMLPFKKSEQEKSLIE 437

Qy      432 KLSMDPAML-----ERLISLDRLASQSGQA-----PILSTPKR 467
Db      438 KVTQLEDLTLLKKEFIQSNKMIKVFREDQIIRLEKL--HKESRGGLPEEQDRLLSELN 495

Qy      468 ERVLMKTVEEKD-----LEIERLKTQKELEAKMLAQKAEKE 506
Db      496 EIOTLREQIEHHPRVAKYAMENHSLRENRRLLEPVPKRAQEMD 540

```

```

RESULT 15
US-09-723-262-2
; Sequence 2, Application US/09723262
; Patent No. 6379912
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,262
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6

```

```

; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
; US-09-723-262-2

Query Match      21.7%; Score 561; DB 3; Length 1388;
Best Local Similarity 30.9%; Pred. No. 5.4e-43;
Matches 162; Conservative 97; Mismatches 192; Indels 74; Gaps 15;

Qy      45 VRVAVRLRPFDV--GTAGASDPPPCVRGMDSCSLEIANRNHQTETKQDFAFYGERSTQ 102
Db      27 IKVVRIRPPAERSGSADGEONLCLSVLSSTSLR---HSPNPKTTFDHDVADVDITQE 83

Qy      103 DIYAGSVQPIRLHLELEGONASVLAYGPTGAGKTHMLGSPREP-----GVIPRA---L 152
Db      84 SVFATVAKSIVSCSMGVNGTIFAYGQTGSKTFTWNGPSESDNFSNLRGVIPRSFEYL 143

Qy      153 MDLLQLTREGEAGRPWALSVTMGLYIYQKVLDDLPASGDLVIREDCRGNLIPGLS 212
Db      144 FSLIDREKEKAGAGKSPFLCK--CSFIEIYNEQIYDLDLSASAGLYLREHIKKGVFVVGAV 201

Qy      213 QKPISSEADFERHFLPASRNRTVGATRLNQRSSSHAVLLVKVDORERLAPP-RQREGKL 271
Db      202 EQVVTSAEAYQVLSGCGWRNRVASTSMNRESSSHAVFTITTESMEKSEINVRTSL 261

Qy      272 YLIDLAGSENRRTGNKGLRKESGAINTSIFVLGKVVDAL---NOGLPRVPYRDSKLT 327
Db      262 NLVDLAGSERQKQTHAEGMRLEKAGNINRSLSCLGVITLVDVGNKGQRHVCYRDSKLT 321

Qy      328 RLQDSLGGSAHSLIANIAPERFYLDTVSALNFAARSKEVINRPTNESLOPHALGPV 387
Db      322 FLRDSLGNAKTAIIANVHPGRCFCGETLSTLNFAQRAKLIKNKAVVNEDTQ-----GNV 377

Qy      388 KLSQ-----KELIG-----PEAKRARGPREE---EIGSPPEMAAPASASOKLSPLQ 431
Db      378 SQLQAEVKRLKEQLAELASGQTPEPESFLTRDKKNTMYEQEAMLPFKKSEQEKSLIE 437

Qy      432 KLSMDPAML-----ERLISLDRLASQSGQA-----PILSTPKR 467
Db      438 KVTQLEDLTLLKKEFIQSNKMIKVFREDQIIRLEKL--HKESRGGLPEEQDRLLSELN 495

Qy      468 ERVLMKTVEEKD-----LEIERLKTQKELEAKMLAQKAEKE 506
Db      496 EIOTLREQIEHHPRVAKYAMENHSLRENRRLLEPVPKRAQEMD 540

```

Search completed: November 5, 2004, 18:46:45  
Job time : 22.0023 secs

This page blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 105.087 Seconds  
(without alignment)  
1747.786 Million cell updates/sec

Title: US-10-797-893-4

Perfect score: 2589

Sequence: 1 MPAAAGSTQORREMAAAS.....LEAKMLQAQAEKENHCPTM 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_23Sep04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	100.0	512	7	ADC23340 Human kin
2	2589	100.0	512	8	ADQ60230 Human mic
3	2577	99.5	665	8	ADQ09240 Human KNS
4	2467	95.3	487	7	ADC23344 Human kin
5	2467	95.3	487	8	ADQ60234 Human mic
6	2419	93.4	490	7	ADK40973 Novel hum
7	2151	83.1	460	3	RAB56650 Human pro
8	1873	72.3	370	7	ADC23338 Human kin
9	1873	72.3	370	8	ADQ60228 Human mic
10	1763	68.1	346	7	ADC23342 Human kin
11	1763	68.1	346	8	ADQ60232 Human mic
12	616	23.8	784	4	ABB71112 Drosophil
13	593	22.9	677	4	ABB65183 Drosophil
14	578.5	22.3	1237	7	ADK40926 Novel hum
15	574.5	22.2	1232	7	ADK48938 Human lun
16	574.5	22.2	1232	7	ADD18924 Human dis
17	574.5	22.2	1232	7	ADJ70347 Human hea
18	574.5	22.2	1235	7	ADK40901 Novel hum
19	573.5	22.2	1234	5	ABG70993 Human HsK
20	568.5	22.0	473	5	ABG70992 Human tar
21	568.5	22.0	522	5	ABG70991 Human HsK
22	568	21.9	898	4	ABU53123 Intracell
23	568	21.9	898	5	Aau76957 Novel hum
24	568	21.9	898	7	ADC31082 Human nov
25	568	21.9	898	7	ADI15915 Human PP

26	567.5	21.9	998	8	ADO44167	Structura
27	564.5	21.8	548	7	ADM04007	Human pro
28	564.5	21.8	905	5	Aau76967	Novel hum
29	561.5	21.7	1232	5	ABG70990	Human HsK
30	561	21.7	1388	5	AAE14400	Human kin
31	561	21.7	1388	5	AAU79590	Human kin
32	561	21.7	1388	6	ABR48222	Human bla
33	561	21.7	1388	7	ADB80468	Ovarian c
34	561	21.7	1388	7	ADC35116	Human bre
35	561	21.7	1388	8	ADL83290	Human PRO
36	561	21.7	1388	8	ADQ20128	Human bof
37	561	21.7	1388	8	ADQ09226	Human KNS
38	560	21.6	972	8	ADN40548	Candida a
39	560	21.6	974	8	ADN40552	Candida a
40	559.5	21.6	1401	7	ADJ94914	Novel NOV
41	557.5	21.5	672	7	ADJ69858	Human hea
42	557.5	21.5	1232	7	ADD49937	Human kin
43	557.5	21.5	1232	7	ADD49932	Human lun
44	557	21.5	1324	7	ADJ95086	Novel NOV
45	556.5	21.5	1066	4	AAG67418	Amino aci

ALIGNMENTS

RESULT 1  
ADC23340  
ID ADC23340 standard; protein; 512 AA.  
XX AC ADC23340;  
XX AC  
DT 18-DEC-2003 (first entry)  
XX  
DE Human kinesin-like DNA binding protein (KID) (SeqID 4).  
XX  
KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "Encoded by CA"  
FT  
XX  
FN US6387644-B1.  
XX  
PD 14-MAY-2002.  
XX  
PF 28-NOV-2000; 2000US-00724224.  
XX  
PR 20-APR-1999; 99US-00295612.  
PR 20-JUN-2000; 2000US-00597292.  
XX  
PA (CYTO-) CYTOKINETICS INC.  
XX  
PI Beraud C;  
XX  
DR WPI; 2003-706919/67.  
DR N-PSDB; ADC23339.  
XX  
PT Identifying a candidate agent as modulator of function of a target  
PT protein for treating cellular proliferation disorders by adding a  
PT candidate agent to a mixture of the target protein that  
PT directly/indirectly produces ADP or phosphate.  
XX  
PS Claim 1; SEQ ID NO 4; 26pp; English.  
XX  
CC This invention relates to a novel method for high throughput screening  
CC systems used to identify compounds for the treatment of cellular  
CC proliferation disorders. Specifically, it refers to candidate agents that  
CC are capable of modulating the activity of target proteins having motor

CC domains, such that the target protein directly or indirectly produces ADP  
CC or phosphate. Furthermore, this activity can be determined using  
CC fluorescence or absorbance readouts. The present invention describes a  
CC method that identifies modulators of the target protein, which is a  
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardinals,  
CC immunomodulators and antiinflammatories. Accordingly, through gene  
CC therapy, they can be used for the treatment of cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
CC polypeptide sequence is human KID protein (SeqID 4) of the invention.  
XX  
XX  
SQ Sequence 512 AA;

Query Match		100.0%;	Score 2589;	DB 7;	Length 512;
Best Local Similarity		100.0%;	Pred. No. 5.2e-233;		
Matches 512;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPAAGSTQORREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60	
DB	1	MPAAGSTQORREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60	
QY	61	ASDPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
DB	61	ASDPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
QY	121	NASVLAYGPTGAGKTHMTLGSPPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEI	180		
DB	121	NASVLAYGPTGAGKTHMTLGSPPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEI	180		
QY	181	YQEKVLDLDPASGDLVIREDCRGNILIPGLSKPTSSFADFERHPLPASRNRVTGATRL	240		
DB	181	YQEKVLDLDPASGDLVIREDCRGNILIPGLSKPTSSFADFERHPLPASRNRVTGATRL	240		
QY	241	NQSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAIN	300		
DB	241	NQSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAIN	300		
QY	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSAHSILANTAPERFVLDTVSAL	360		
DB	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSAHSILANTAPERFVLDTVSAL	360		
QY	361	NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAP	420		
DB	361	NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAP	420		
QY	421	ASASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMKTVEEKD	480		
DB	421	ASASQKLSPLQKSSMDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMKTVEEKD	480		
QY	481	LEIERLTKQKELEAKMLAKAEKEKNCPTM	512		
DB	481	LEIERLTKQKELEAKMLAKAEKEKNCPTM	512		

RESULT 2  
ADQ60230  
ID ADQ60230 standard; protein; 512 AA.  
XX  
XX AC ADQ60230;  
XX  
XX DT 23-SEP-2004 (first entry)  
XX  
XX DE Human microtubule motor protein #2.  
XX  
XX KW Human; microtubule motor protein; cellular proliferation disorder;  
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
XX  
XX OS Homo sapiens.  
XX  
XX PN US6762043-B1.  
XX  
XX PD 13-JUL-2004.

XX  
PF 06-MAR-2002; 2002US-00093317.  
XX  
XX 20-APR-1999; 99US-00295612.  
PR 20-JUN-2000; 2000US-00597292.  
PR 28-NOV-2000; 2000US-00724224.  
XX  
XX (CYTO-) CYTOKINETICS INC.  
PA  
XX Beraud C;  
XX  
XX WPI; 2004-532491/51.  
XX  
XX New isolated microtubule motor protein, useful for screening modulators  
PT for treating cellular proliferation disorders such as cancer,  
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
PT inflammation.  
XX  
XX Claim 1; SEQ ID NO 4; 26pp; English.  
XX  
XX The invention relates to human microtubule motor proteins and the nucleic  
CC acids encoding them. The invention also relates to a method of screening  
CC for modulators of a motor protein which has microtubule stimulated ATPase  
CC activity, a method of testing for ATPase activity of microtubule motor  
CC proteins, methods to identify candidate agents that bind to a target  
CC protein or act as a modulator of the binding characteristics or  
CC biological activity of a target protein, modulators of the target  
CC protein, and methods of treating cellular proliferation disorders such as  
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
CC and inflammation, for treating disorders associated with kinesin-like DNA  
CC binding protein (KID) and for inhibiting KID. The sequences are used for  
CC screening for modulators of motor proteins useful for treating cellular  
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
CC hypertrophy, immune disorders and inflammation, for treating disorders  
CC associated with KID and for inhibiting KID and for treating autoimmune  
CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents a human microtubule motor protein  
CC of the invention. Note: The specification states that this sequence is  
CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not  
CC appear to be the case.  
XX  
XX SQ Sequence 512 AA;

Query Match		100.0%;	Score 2589;	DB 8;	Length 512;
Best Local Similarity		100.0%;	Pred. No. 5.2e-233;		
Matches 512;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MPAAGSTQORREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60	
DB	1	MPAAGSTQORREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRPFVDGTAG	60	
QY	61	ASDPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
DB	61	ASDPPCVRGMDSCSLEIANWRNHQETLKYQDAFYGERSTQODIYAGSVQPIRLHLEQ	120		
QY	121	NASVLAYGPTGAGKTHMTLGSPPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEI	180		
DB	121	NASVLAYGPTGAGKTHMTLGSPPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEI	180		
QY	181	YQEKVLDLDPASGDLVIREDCRGNILIPGLSKPTSSFADFERHPLPASRNRVTGATRL	240		
DB	181	YQEKVLDLDPASGDLVIREDCRGNILIPGLSKPTSSFADFERHPLPASRNRVTGATRL	240		
QY	241	NQSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAIN	300		
DB	241	NQSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAIN	300		
QY	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSAHSILANTAPERFVLDTVSAL	360		
DB	301	SLFVLGKVDALNQGLPRVYRDSKLTLLQDSLGSAHSILANTAPERFVLDTVSAL	360		
QY	361	NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAP	420		





CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention.

XX SQ Sequence 487 AA;

Query Match 95.3%; Score 2467; DB 8; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-221;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GCRLSKIGATRRPPARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 86  
 DB 2 GCRLSKIGATRRPPARVRVAVLRPFVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 QY 87 LKYPDAFYGERSTQDDIYAGSVQPIRLHLLGGQNASVLAYGPTGAGKTHMLGSPQPG 146  
 DB 62 LKYPDAFYGERSTQDDIYAGSVQPIRLHLLGGQNASVLAYGPTGAGKTHMLGSPQPG 121  
 QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRNI 206  
 DB 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRNI 181  
 QY 207 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQSRSSHAVLLVKVDQERLAPRQ 266  
 DB 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQSRSSHAVLLVKVDQERLAPRQ 241  
 QY 267 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVDALNQLPRVPYRDSKL 326  
 DB 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVDALNQLPRVPYRDSKL 301  
 QY 327 TRLQDSLGGSAHSILIANIAPRRFYLDTVSALNFAARSKEVINRPTNESLQPHALGP 386  
 DB 302 TRLQDSLGGSAHSILIANIAPRRFYLDTVSALNFAARSKEVINRPTNESLQPHALGP 361  
 QY 387 VKLSQKELGPPKARAGPPEERTGSPPEWAAASQKLSPLQKSSMDPAMLERLLS 446  
 DB 362 VKLSQKELGPPKARAGPPEERTGSPPEWAAASQKLSPLQKSSMDPAMLERLLS 421  
 QY 447 LDRLASQSGQAPILSTPKERVMVMTVEKOLIEIRLTKQKLEAKMLAKQAEBKE 506  
 DB 422 LDRLASQSGQAPILSTPKERVMVMTVEKOLIEIRLTKQKLEAKMLAKQAEBKE 481  
 QY 507 NHCPTM 512  
 DB 482 NHCPTM 487

## RESULT 6

ADK40973  
 ID ADK40973 standard; protein; 490 AA.

XX AC ADK40973;

XX DT 06-MAY-2004 (first entry)

XX DE Novel human kinase protein #80.

XX cytosolic; immunomodulator; cardiant; neuroprotective; nootropic;  
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
 KW cancer; peripheral nervous system; central nervous system;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; viral infection; prion infection;  
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
 KW attention disorder; cognitive disorder; hypotension; hypertension;  
 KW psychotic disorder; neurological disorder; dyskinesia;  
 KW metabolic disorder; organ transplant rejection; enzyme.

OS Homo sapiens.

XX WO2003057841-A2.

XX PD 17-JUL-2003.

XX PF 31-DEC-2002; 2002WO-US041687.

XX PR 31-DEC-2001; 2001US-0343169P.

XX PA (GRIG/) GRIGORIEV I V.

XX PI (SUDA/) SUDARSANAM S.

XX PT Grigoriev IV, Sudarsanam S;

XX WPI; 2003-587115/55.

XX New isolated, enriched or purified nucleic acid molecule encoding a  
 PT kinase polypeptide, useful for treating cancer, immune-related diseases,  
 PT cardiovascular disease, brain or neuronal-associated diseases and  
 PT metabolic disorders.

PS Claim 1; SEQ ID NO 80; 491pp; English.

XX The invention relates to novel isolated, enriched or purified nucleic acid  
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule  
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and  
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the  
 CC polypeptide in (a), except that it lacks one or more, but not all, of an  
 CC N-terminal domain, C-terminal catalytic domain, a catalytic region and a C-  
 CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,  
 CC polypeptides, methods and substance are useful for treating cancers,  
 CC immune-related diseases or disorders, cardiovascular disease, brain or  
 CC neuronal-associated diseases, and metabolic disorders. The disorders are  
 CC preferably cancers of the tissues or of hematopoietic origin, diseases of  
 CC the central or peripheral nervous system, Alzheimer's disease,  
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,  
 CC viral infections, infections caused by prions, infections caused by  
 CC bacteria, infections caused by fungi, ocular diseases, migraines, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognition  
 CC disorders, hypotension, hypertension, psychotic disorders, neurological  
 CC disorders, dyskinesias, metabolic disorders and organ transplant  
 CC rejection. This sequence corresponds to one of the kinase polypeptides of  
 CC the invention.

XX SQ Sequence 490 AA;

Query Match 93.4%; Score 2419; DB 7; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-217;  
 Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGSTQORREMAAASAAATISGAGRCRLSKIGATRRPPPARVRVAVLRPFVDTAGAS 62

DB 10 AAGSTQORREMAAASAAATISGAGRCRLSKIGATRRPPPARVRVAVLRPFVDTAGAS 69

QY 63 DPCVCGMDSCLSIANRNHOETLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLGGQNA 122

DB 70 DPCVCGMDSCLSIANRNHOETLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLGGQNA 129

QY 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182

DB 130 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 189

QY 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQ 242

DB 190 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQ 249

QY 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAINTS 302

DB 250 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRRTGNKGLRKESGAINTS 309

QY 303 FVLGVVDALNOGLPRVPRVDSKLTLLQSLGSSAHSILIANIAPERRFYLDTVSALNF 362  
DB 310 FVLGVVDALNOGLPRVPRVDSKLTLLQSLGSSAHSILIANIAPERRFYLDTVSALNF 369  
QY 363 AARSKEVINPFTNESLOPHALGPVKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPAS 422  
DB 370 AARSKEVINPFTNESLOPHALGPVKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPAS 429  
QY 423 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKOL 481  
DB 430 ASQKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKOL 488

RESULT 7  
AAB56650  
ID AAB56650 standard; protein; 460 AA.  
XX AAB56650;  
AC AAB56650;  
DT 13-MAR-2001 (first entry)  
XX Human prostate cancer antigen protein sequence SEQ ID NO:1228.  
DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
XX neuroprotective; cytostatic; cardiocactive; immunomodulatory; muscular;  
XX vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
XX antibacterial; gene therapy; immune; reproductive; renal;  
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
XX wound; infectious disease.  
XX Homo sapiens.  
OS Homo sapiens.  
XX WO200055174-A1.  
PN 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US005988.  
PF 12-MAR-1999; 99US-0124270P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX Rosen CA, Ruben SM;  
FI WPI; 2000-587513/55.  
XX N-PSDB; AAF15853.  
DR Prostate cancer associated gene sequences, referred to as prostate cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as prostate cancer.  
XX Claim 11; Page 1649-1651; 2338pp; English.  
XX AAF15566 to AAF16505 encode the human prostate cancer associated  
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
CC The prostate cancer antigens can have neuroprotective, cytostatic,  
CC cardiocactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAB57303 represent sequences used in the exemplification of the present  
CC invention  
XX Sequence 460 AA;

Query Match 83.1%; Score 2151; DB 3; Length 460;  
Best Local Similarity 95.9%; Pred. No. 4.6e-192;

Matches 424; Conservative 1; Mismatches 17; Indels 0; Gaps 0;  
QY 25 GAGRCRLSKI GATRRPPPPARVRVRLPPFVDTAGASDPPCVRGMDSCSLEIANWRNHQ 84  
DB 2 GAGRCRLSKI GATRRPPPPARVRVRLPPFVDTAGASDPPCVRGMDSCSLEIANWRNHQ 61  
QY 85 ETLKYQDFAPYGERSTOODIYAGSVQPIRLRHLLGQNASVLAYGPTGAGKTHMLGSPEQ 144  
DB 62 ETLKYQDFAPYGERSTOODIYAGSVQPIRLRHLLGQNASVLAYGPTGAGKTHMLGSPEQ 121  
QY 145 PGVIPRALMDLLQLTRREGAERGPAWLSVTWMSYLEIYQEKVLDLDDPASGDLVIREDCRG 204  
DB 122 PGVIPRALMDLLQLTRREGAERGPAWLSVTWMSYLEIYQEKVLDLDDPASGDLVIREDCRG 181  
QY 205 NILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLAPF 264  
DB 182 NILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLAPF 241  
QY 265 ROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVVDALNOGLPRVPRYRDS 324  
DB 242 ROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGVVDALNOGLPRVPRYRDS 301  
QY 325 KLTRLQLQSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVINPFTNESLOPHAL 384  
DB 302 KLTRLQLQSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVINPFTNESLOPHAL 361  
QY 385 GPVKLSQKELGPPPEAKRARGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLERL 444  
DB 362 GPVKLSQKELGPPPEAKRARGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLERL 421  
QY 445 LSLDRLLASQSGQAPLLSTPK 466  
DB 422 LQGPSACLPGEPXGFSVEYPK 443

RESULT 8  
ADC23338  
ID ADC23338 standard; protein; 370 AA.  
XX ADC23338;  
AC ADC23338;  
DT 18-DEC-2003 (first entry)  
XX Human kinesin-like DNA binding protein (KID) (SeqID 2).  
DE human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
XX cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
XX cancer; hyperplasia; restenosis; cellular proliferation disorder;  
XX cardiac hypertrophy; immune disorder; inflammation.  
XX Homo sapiens.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FT Misc-difference 1 /note= "Encoded by ATGCA"  
FT US6387644-B1.  
XX PN 14-MAY-2002.  
XX PD 28-NOV-2000; 2000US-0072422A.  
XX PF 20-APR-1999; 99US-00295612.  
XX PR 20-JUN-2000; 2000US-00597292.  
XX PX (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C;  
XX WPI; 2003-706919/67.  
XX N-PSDB; ADC23337.  
XX Identifying a candidate agent as modulator of function of a target  
PT

PT protein for treating cellular proliferation disorders by adding a  
PT candidate agent to a mixture of the target protein that  
PT directly/indirectly produces ADP or phosphate.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC This invention relates to a novel method for high throughput screening  
XX systems used to identify compounds for the treatment of cellular  
XX proliferation disorders. Specifically, it refers to candidate agents that  
XX are capable of modulating the activity of target proteins having motor  
XX domains, such that the target protein directly or indirectly produces ADP  
XX or phosphate. Furthermore, this activity can be determined using  
XX fluorescence or absorbance readouts. The present invention describes a  
XX method that identifies modulators of the target protein, which is a  
XX kinesin-like DNA binding protein (known as KID) as cytosolic, cardinals,  
XX immunomodulators and antiinflammatory. Accordingly, through gene  
XX therapy, they can be used for the treatment of cancer, hyperplasias,  
XX restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
XX polypeptide sequence is human KID protein (SeqID 2) of the invention.

XX SQ Sequence 370 AA;

Query Match 72.3%; Score 1873; DB 7; Length 370;

Best Local Similarity 100.0%; Pred. No. 3.5e-166; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGSTQORREMAAASAAAGAGRCRLSKIGATRRPPARVAVRLRPFVDTAGAS 62  
Db 2 AAGSTQORREMAAASAAAGAGRCRLSKIGATRRPPARVAVRLRPFVDTAGAS 61  
Qy 63 DPPCVRGMDSCSLEIANWRNHQETLKQFDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNA 122  
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKQFDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNA 121  
Qy 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREAGRPGWALSVMYSLEYQ 182  
Db 122 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREAGRPGWALSVMYSLEYQ 181  
Qy 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 242  
Db 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 241  
Qy 243 RSSRSHAVLLVKVDQERLAPFRQREGKLYIDLAGEDNRRRTGNKGLRKESGAINITSL 302  
Db 242 RSSRSHAVLLVKVDQERLAPFRQREGKLYIDLAGEDNRRRTGNKGLRKESGAINITSL 301  
Qy 303 FVLGKVVVDALNQGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 362  
Db 302 FVLGKVVVDALNQGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 361  
Qy 363 AARSKEVIN 371  
Db 362 AARSKEVIN 370

RESULT 9

ID ADQ60228 standard; protein; 370 AA.

XX AC ADQ60228;

XX DT 23-SEP-2004 (first entry)

XX DE Human microtubule motor protein #1.

KW Human; microtubule motor protein; cellular proliferation disorder;  
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

OS Homo sapiens.

XX US6762043-B1.

PN

XX 13-JUL-2004.

XX 06-MAR-2002; 2002US-00093317.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX 28-NOV-2000; 2000US-00724224.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2004-532491/51.

XX New isolated microtubule motor protein, useful for screening modulators  
XX for treating cellular proliferation disorders such as cancer,  
XX hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
XX inflammation.

XX Claim 1; SEQ ID NO 2; 26pp; English.

XX CC The invention relates to human microtubule motor proteins and the nucleic  
XX acids encoding them. The invention also relates to a method of screening  
XX for modulators of a motor protein which has microtubule stimulated ATPase  
XX activity, a method of testing for ATPase activity of microtubule motor  
XX proteins, methods to identify candidate agents that bind to a target  
XX protein or act as a modulator of the binding characteristics or  
XX biological activity of a target protein, modulators of the target  
XX protein, and methods of treating cellular proliferation disorders such as  
XX cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
XX and inflammation, for treating disorders associated with kinesin-like DNA  
XX binding protein (KID) and for inhibiting KID. The sequences are used for  
XX screening for modulators of motor proteins useful for treating cellular  
XX proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
XX hypertrophy, immune disorders and inflammation, for treating disorders  
XX associated with KID and for inhibiting KID and for treating autoimmune  
XX diseases, arthritis, graft rejection, inflammatory bowel disease and  
XX angioplasty. This sequence represents a human microtubule motor protein  
XX of the invention. Note: The specification states that this sequence is  
XX encoded by the nucleic acid featured as SEQ ID NO:1, but this does not  
XX appear to be the case.

XX SQ Sequence 370 AA;

Query Match 72.3%; Score 1873; DB 8; Length 370;

Best Local Similarity 100.0%; Pred. No. 3.5e-166;

Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGSTQORREMAAASAAAGAGRCRLSKIGATRRPPARVAVRLRPFVDTAGAS 62  
Db 2 AAGSTQORREMAAASAAAGAGRCRLSKIGATRRPPARVAVRLRPFVDTAGAS 61  
Qy 63 DPPCVRGMDSCSLEIANWRNHQETLKQFDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNA 122  
Db 62 DPPCVRGMDSCSLEIANWRNHQETLKQFDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNA 121  
Qy 123 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREAGRPGWALSVMYSLEYQ 182  
Db 122 SVLAYGPTGAGKTHMLGSPQPGVIPRALMDLLQLTREAGRPGWALSVMYSLEYQ 181  
Qy 183 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 242  
Db 182 EKVLDDLPASGDLVIREDCRGNILIPGLSQKPISSFADFRRHFLPASRNTVATRLNQ 241  
Qy 243 RSSRSHAVLLVKVDQERLAPFRQREGKLYIDLAGEDNRRRTGNKGLRKESGAINITSL 302  
Db 242 RSSRSHAVLLVKVDQERLAPFRQREGKLYIDLAGEDNRRRTGNKGLRKESGAINITSL 301  
Qy 303 FVLGKVVVDALNQGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 362  
Db 302 FVLGKVVVDALNQGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERRFYLDTVSALNF 361

QY 363 AARSKEVIN 371  
Db 362 AARSKEVIN 370

RESULT 10  
ADC23342  
ID ADC23342 standard; protein; 346 AA.  
XX AC ADC23342;  
XX DT 18-DEC-2003 (first entry)  
XX DE Human kinesin-like DNA binding protein (KID) (SeqID 6).  
XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX OS Homo sapiens.  
XX FN US6387644-B1.  
XX PD 14-MAY-2002.  
XX PF 28-NOV-2000; 2000US-00724224.  
XX PR 20-APR-1999; 99US-00295612.  
XX PR 20-JUN-2000; 2000US-00597292.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C;  
XX WPI; 2003-706919/67.  
XX DR N-PSDB; ADC23341.  
XX PT Identifying a candidate agent as modulator of function of a target  
XX PT protein for treating cellular proliferation disorders by adding a  
XX PT candidate agent to a mixture of the target protein that  
XX PT directly/indirectly produces ADP or phosphate.  
XX PS Claim 1; SEQ ID NO 6; 26pp; English.  
XX SQ

This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiant, immunomodulators and antiinflammatory. Accordingly, through gene therapy, they can be used for the treatment of cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation. This polypeptide sequence is human KID protein (SeqID 6) of the invention.

Query Match 68.1%; Score 1763; DB 7; Length 346;  
Best Local Similarity 100.0%; Pred. No. 6.4e-156;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCRLSKIGATRRPPPARVAVRLRPFDGTAGADPPCVRGMDSCSLEIANWRNHQET 86  
Db 2 GRCRLSKIGATRRPPPARVAVRLRPFDGTAGADPPCVRGMDSCSLEIANWRNHQET 61

QY 87 LKYQFDAYGERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTCAGKTHMLGSPQPG 146  
Db 62 LKYQFDAYGERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTCAGKTHMLGSPQPG 121

QY 147 VIPRALMDLLQLTREAGRPNWALSVTMSYLIYQEKVLDLDPASGDLVIREDCGNI 206  
Db 122 VIPRALMDLLQLTREAGRPNWALSVTMSYLIYQEKVLDLDPASGDLVIREDCGNI 181

QY 207 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPFRQ 266  
Db 182 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPFRQ 241

QY 267 REGKLYLIDLAGSEDDNRRTGNKGLRLKESGAINTSFLVLGKVVQDALNQGLPRPYRDSKL 326  
Db 242 REGKLYLIDLAGSEDDNRRTGNKGLRLKESGAINTSFLVLGKVVQDALNQGLPRPYRDSKL 301

QY 327 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 371  
Db 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 11  
ADQ60232  
ID ADQ60232 standard; protein; 346 AA.  
XX AC ADQ60232;  
XX DT 23-SEP-2004 (first entry)  
XX DE Human microtubule motor protein #3.  
XX KW Human; microtubule motor protein; cellular proliferation disorder;  
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
XX OS Homo sapiens.  
XX FN US6762043-B1.  
XX PD 13-JUL-2004.  
XX PF 06-MAR-2002; 2002US-00093317.  
XX PR 20-APR-1999; 99US-00295612.  
XX PR 20-JUN-2000; 2000US-00597292.  
XX PR 28-NOV-2000; 2000US-00724224.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C;  
XX WPI; 2004-532491/51.  
XX DR N-PSDB; ADQ60231.  
XX PT New isolated microtubule motor protein, useful for screening modulators  
XX PT for treating cellular proliferation disorders such as cancer,  
XX PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
XX PT inflammation.  
XX PS Claim 1; SEQ ID NO 6; 26pp; English.  
XX SQ

The invention relates to human microtubule motor proteins and the nucleic acids encoding them. The invention also relates to a method of screening for modulators of a motor protein which has microtubule stimulated ATPase activity, a method of testing for ATPase activity of microtubule motor proteins, methods to identify candidate agents that bind to a target protein or act as a modulator of the binding characteristics or biological activity of a target protein, modulators of the target protein, and methods of treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with kinesin-like DNA binding protein (KID) and for inhibiting KID. The sequences are used for screening for modulators of motor proteins useful for treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders



CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention.  
 XX  
 SQ Sequence 346 AA;

Query Match 68.1%; Score 1763; DB 8; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-156;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 27 GRCRLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86  
 DB 2 GRCRLSKIGATRRPPPARVRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 QY 87 LKYQFDAYGERSQODIYAGSVQPIILRHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 146  
 DB 62 LKYQFDAYGERSQODIYAGSVQPIILRHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121  
 QY 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLLOPASGDVIREDCRGN 206  
 DB 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLLOPASGDVIREDCRGN 181  
 QY 207 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSRAVLLVKVDQRLAPPRQ 266  
 DB 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSRAVLLVKVDQRLAPPRQ 241  
 QY 267 REGKYLIDLAGSEDRNRGTGKRLKESGAINTSFLVGLKVDALNQLPRVYRDSKL 326  
 DB 242 REGKYLIDLAGSEDRNRGTGKRLKESGAINTSFLVGLKVDALNQLPRVYRDSKL 301  
 QY 327 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 371  
 DB 302 TRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346

RESULT 12  
 ABB71112  
 ID ABB71112 standard; protein; 784 AA.  
 AC ABB71112;  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX  
 XX Drosophila melanogaster polypeptide SEQ ID NO 40128.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 XX Drosophila melanogaster.  
 XX  
 XX WO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 XX  
 XX 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 XX WPI; 2001-656860/75.  
 XX  
 XX N-PSDB; ABL15215.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signaling and cell-cell  
 XX interactions.  
 XX  
 XX Disclosure; SEQ ID NO 40128; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 784 AA;

Query Match 23.8%; Score 616; DB 4; Length 784;  
 Best Local Similarity 33.1%; Pred. No. 6.1e-48;  
 Matches 176; Conservative 84; Mismatches 183; Indels 88; Gaps 18;  
 QY 37 TRRP-----PPARVRVAVRLRPFDGTAGASDPPCV-----RGMDSCSLEIANW--RN 82  
 DB 5 SRRFGTSSQTFNECVVVRCPMSNRERSERSPEVVVYVNRGV-----VELQNVVDGN 60  
 QY 83 HQETLKQFDAYGERSQODIYAGSVQPIILRHLLLEGQNASVLYAGPTGAGKTHM---L 139  
 DB 61 KEQRKVFTYDAAYDASATQTLTYHEVFPVSVLEGFNGCIFYAGTGTGKTFTMEGVR 120  
 QY 140 GSPEQPGVPRALMDL-LQLTREEGAGRPWALSVTMSYLEIQEKVLDLLOPASGDV 198  
 DB 121 GNDELGMIIPTFQIWLHNRNTE-----FQFLVDVSYLEIYMEELRDLKPNKSHLEV 175  
 QY 199 REDCRGN-ILIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQRSSRSRAVLLVKVDQ 257  
 DB 176 RE--RGSVVVFNHAINCKSVEDMIKVMQVGNKRTVGTFTMNEHSSRSRAIPMIKTEM 233  
 QY 258 RERLAPPRQREGKYLIDLAGSEDRNRGTGKRLKESGAINTSFLVGLKVDALNQLP 317  
 DB 234 CD-TETNTIKVGNLIDLAGSERQSKTGAERLKEASKINLALSSLGNVISALAESPP 292  
 QY 318 RVPYRDSKLTRLLQDSLGGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPF 377  
 DB 293 HVPYRDSKLTRLLQDSLGGSKTMIANIGSNYNYNETLTTRYAGRAKSIQNPQIKNE 352  
 QY 378 SLQPHALGPVKLSQ-----KELGPPPAKRA-----RGPEEEI----- 411  
 DB 353 DPQ-----DAKLKEYQEEIERLKLKELIGPQQQORSEKQVAKKQVKKPKETVTKEMSDS 407  
 QY 412 -----GSPPEMAAPASQKLSPLQKLSMDPAMLERLISLDRLLASQGSQG 458  
 DB 408 LQVSTIEQPVEDSDPE---GAESSEDKENAEVAKSNEELERERVEN-SKLAALAELE 463  
 QY 459 APLLSTPKRERVMKTVKEKLEIER-----LTKQKELEAKMAQKAE 504  
 DB 464 GQLVRGGKN-----LLDTYSEKQIELEKLVIAERKKREIEIQOOLEQEE 510

RESULT 13  
 ABB65183  
 ID ABB65183 standard; protein; 677 AA.  
 XX  
 XX ABB65183;  
 XX  
 XX 26-MAR-2002 (first entry)  
 XX  
 XX Drosophila melanogaster polypeptide SEQ ID NO 22341.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 XX Drosophila melanogaster.  
 XX  
 XX WO200171042-A2.  
 XX  
 XX 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PR 11-JUL-2000; 2000US-00614150.  
XX PA (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX MPI; 2001-656860/75.  
XX N-PSDB; ABL09286.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX Disclosure; SEQ ID NO 22341; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
XX ABB72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 677 AA;  
Query Match 22.9%; Score 593; DB 4; Length 677;  
Best Local Similarity 33.9%; Pred. No. 6.9e-46;  
Matches 167; Conservative 79; Mismatches 190; Indels 58; Gaps 12;  
QY 45 VRVAVRLRPFDG--TAGADPCVCGMD--SCSLEIANWNRHQTLYQDFAFGERSTQ 101  
DB 21 VRVVRTRPMDKNELSAGALSIVDKINRAITVMKPNATANEPPKTYTFDNNVFDGGSNQ 80  
QY 102 QDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGTHTMLGSPEOP---GVIPRALMDL--- 155  
DB 81 MDLYVDTRARIVKVLGYNGTILAYGQTGTGTYTMSGNPDSPQTKGIIPNAFAHIFGH 140  
QY 156 LQITREAGRGPRWALSVTWMSYLEIYQEKVLDLLDPASG--DLVIREDCRGNILIPGLSQK 214  
DB 141 IAKAKEN-----QKFLVRVSYMEIYNEEVRDLLGKDVGKSLVKRERPDIGVFVKLSGY 194  
QY 215 PISFADFERHFLPASNRRTVGATRLNQRSSRHAVLLVKVDORERLAPRQ--REGKLY 272  
DB 195 VVHNADLENIMRLGNKRVAGATKMNQESRSHAFITVSELSGEGDVQVVRMGKLLQ 254  
QY 273 LIDLAGSDNRNTGNKGLRKESGAINTSILVLGKVVADLNQ--LPRVPVYRDSKLTLLQ 331  
DB 255 LVDLAGSERQSKTQASQRLKEATKINLSVLGNVISALVDGKSTHIPPYNSKLTLLQ 314  
QY 332 DSLGGSASHILIANIAPERFYLDTVSALNFAARKEVINRPTNE-----SLQPHALGPV 387  
DB 315 DSLGNSKTVCATISPADSNYMETITSLRYASRAKNIQNRMEINEEPKDALLRHFQEEI 374  
QY 388 KLSQKEL-LGPPEAKRARGPEERIGSPPEWAPASASQKLSPLQKLSMDPAMLERLLS 446  
DB 375 ARLRKQLEGGUSLEERPPSEEREDTADDELEAPLELESSSTIOAVEKPKKREK--- 431  
QY 447 LDRLLASQGSQAGPFLSTPKRMRVMVMTVEEKDLEIERLTKQKLEAKM----- 497  
DB 432 -----TDAEKELAKRNEHQEIEHAKTEQETLRNKLVSLEGIKLV 473  
QY 498 -----LAQKAEKE 506  
DB 474 GGENLEKAQTOQE 486

RESULT 14  
ADK40926  
XX ID ADK40926 standard; protein; 1237 AA.  
XX AC ADK40926;  
XX DT 06-MAY-2004 (first entry)  
XX DE Novel human kinase protein #33.  
XX cytotactic; immunomodulator; cardiant; neuroprotective; nootropic;  
XX antiparkinsonian; virucide; antibacterial; kinase inhibitor; kinase;  
XX analgesic; hypotensive; immunosuppressive; central nervous system;  
XX cancer; peripheral nervous system; multiple sclerosis;  
XX Alzheimer's disease; Parkinson's disease; viral infection; prion infection;  
XX amyotrophic lateral sclerosis; pain; sexual dysfunction; mood disorder;  
XX ocular disease; migraine; cognition disorder; hypotension; hypertension;  
XX attention disorder; neurological disorder; dyskinesia;  
XX psychototic disorder; organ transplant rejection; enzyme.  
XX Homo sapiens.  
XX OS WO2003057841-A2.  
XX PN 17-JUL-2003.  
XX PD 31-DEC-2002; 2002WO-US041687.  
XX PF 31-DEC-2001; 2001US-0343169P.  
XX PR (GRIG/) GRIGORIEV I V.  
XX PA (SUDA/) SUDARSANAM S.  
XX PI Grigoriev IV, Sudarsanam S;  
XX WPI; 2003-587115/55.  
XX New isolated, enriched or purified nucleic acid molecule encoding a  
XX kinase polypeptide, useful for treating cancer, immune-related diseases,  
XX cardiovascular disease, brain or neuronal-associated diseases and  
XX metabolic disorders.  
XX Claim 1; SEQ ID NO 33; 491pp; English.  
XX The invention relates to novel isolated, enriched or purified nucleic acid  
XX molecules encoding a kinase polypeptide. The nucleic acid molecule  
XX comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
XX complement of (a); (c) hybridizes under stringent conditions to (a) and  
XX encodes a naturally occurring kinase polypeptide; (d) encodes the  
XX polypeptide in (a), except that it lacks one or more, but not all, of an  
XX N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-  
XX terminal domain, a coiled-coil structure region, a spacer region and a C-  
XX terminal tail; or (e) is a complement of (d). The nucleic acid molecules,  
XX polypeptides, methods and substance are useful for treating cancers,  
XX immune-related diseases or disorders, cardiovascular disease, brain or  
XX neuronal-associated diseases, and metabolic disorders. The disorders are  
XX preferably cancers of the tissues or of hematopoietic origin, diseases of  
XX the central or peripheral nervous system, Alzheimer's disease,  
XX Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,  
XX viral infections, infections caused by prions, infections caused by  
XX bacteria, infections caused by fungi, ocular diseases, migraines, pain,  
XX sexual dysfunction, mood disorders, attention disorders, cognitive  
XX disorders, hypotension, hypertension, psychototic disorders, neurological  
XX disorders, dyskinesias, metabolic disorders and organ transplant  
XX rejection. This sequence corresponds to one of the kinase polypeptides of  
XX the invention.  
XX Sequence 1237 AA;  
Query Match 22.3%; Score 578.5; DB 7; Length 1237;  
Best Local Similarity 32.4%; Pred. No. 4.1e-44;  
Matches 157; Conservative 97; Mismatches 172; Indels 59; Gaps 13;



Db	286	VPYRDKLRLQLQDSLGGNSHTLMIACVSPADSNLEETLNTLRYADRAKKIKNKPVIMID	345
Qy	379	LQPHALGPVK--LSQKELLGPPPEAKRGPEEEIGSPFMAAPASASOKLSP-LQKLSS	435
Db	346	PQPAELNHLKQVQQLQVL-----LLOAHGG-----TLFGSITVEPSENLSLME	390
Qy	436	MDPAMLERLLSLDRLLASQSGCAPLLSTPKRFRMVLMTVEEK-DLEIERLKTOK-EL	493
Db	391	KNQSLVEENEKLSRGLSEAGQTAQML-----ERILTEQANEKNNAKLEELRQHAACKL	445
Qy	494	EAKMLAQKAEKE	506
Db	446	DLQKLVELLEDQE	458

Search completed: November 5, 2004, 18:44:39  
Job time : 107.087 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 10.7872 Seconds  
(without alignments)  
3300.235 Million cell updates/sec

Title: US-10-797-893-2  
Perfect score: 1878  
Sequence: 1 MAAGSTQORREMAAASAA.....FYLDVTVALNFAARSKEVIN 370

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1665	88.7	665	S62328	kinesin-like DNA b
2	594.5	31.7	664	T48258	kinesin-like prote
3	555.5	29.6	784	A55236	kinesin-related pr
4	550	29.3	703	A86319	F25116.11 protein
5	547	29.1	1254	T18277	kinesin heavy chai
6	536.5	28.6	1058	T47525	kinesin-related pr
7	536	28.5	784	T40594	probable kinesin-1
8	532	28.3	1226	I51617	kinesin-like prote
9	531.5	28.3	885	D86151	F22W8.8 protein -
10	527.5	28.1	1388	T30335	KLP2 protein - Afr
11	525	28.0	1070	T06733	kinesin homolog F2
12	524	27.9	1056	H84777	probable kinesin-r
13	521.5	27.8	1006	T02017	kinesin-related pr
14	517.5	27.6	1076	B84687	probable kinesin-1
15	514.5	27.4	699	S38982	kinesin-related pr
16	514	27.4	1066	I A48669	kinesin-related pr
17	514	27.4	1695	A56921	kinesin family pro
18	513.5	27.3	701	B44259	kinesin-related pr
19	513.5	27.3	805	S64238	kinesin-related pr
20	513.5	27.3	1225	A56514	chromokinesin - ch
21	513	27.3	747	A57107	kinesin-related pr
22	511.5	27.2	1022	E84792	probable kinesin h
23	509.5	27.1	786	A53939	kinesin homolog KH
24	509.5	27.1	2954	T14156	kinesin-related pr
25	508	27.1	1229	T48959	kinesin-like prote
26	507.5	27.0	1231	A54803	microtubule-associ
27	507	27.0	928	T10164	kinesin heavy chai
28	504.5	26.9	330	B48835	kinesin-like prote
29	504	26.8	813	T46242	kinesin-like prote

30	501	26.7	1056	1	G02157	kinesin-like epind
31	500	26.6	1150	1	A55289	kinesin-like prote
32	498.5	26.5	935	2	T51930	kinesin [imported]
33	495.5	26.4	958	2	T20621	hypothetical prote
34	495	26.4	929	2	T51932	kinesin [imported]
35	494.5	26.3	883	2	T40128	kinesin-like prote
36	491	26.1	963	1	A41919	kinesin heavy chai
37	490	26.1	742	1	S58691	kinesin-related pr
38	488.5	26.0	1085	2	T38378	kinesin-like prote
39	488	26.0	987	2	T51360	kinesin-like heavy
40	485.5	25.9	793	2	JCS831	kinesin-related pr
41	485.5	25.9	843	2	S44868	kinesin heavy chai
42	483	25.7	332	2	C48835	kinesin-like prote
43	482.5	25.7	1584	1	JN0114	kinesin-related pr
44	482.5	25.7	1584	2	T15822	kinesin-like prote
45	482	25.7	1027	2	S37711	kinesin heavy chai

ALIGNMENTS

RESULT 1

S62328  
kinesin-like DNA binding protein KID - human  
C:Species: Homo sapiens (man)  
C>Date: 23-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 16-Aug-2004  
C:Accession: S62328  
R:Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Tsukita, S.; Inoue, J.  
EMBO J. 15, 457-467, 1996  
A:Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and  
A:Reference number: S62328; MUID:96174806; PMID:8599929  
A:Accession: S62328  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-665 <TK>  
C:Superfamily: kinesin motor domain homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:44-374/Domain: kinesin motor domain homology <KMOT>  
F:128-135/Region: nucleotide-binding motif A (P-loop)

Query Match 88.7%; Score 1665; DB 2; Length 665;  
Best Local Similarity 91.4%; Pred. No. 2.4e-125;  
Matches 339; Conservative 1; Mismatches 29; Indels 2; Gaps 2;

Qy	1	MAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPPARVAVLRFPVDTGAGA	60
Db	1	MAAGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPPARVAVLRFPVDTGAGA	60
Qy	61	SDPPCVRGMDSCSLEIANRNHQETLKYQDAFYGERSTQODIYAGSVQPIRLHLEGGN	120
Db	61	SDPPCVRGMDSCSLEIANRNHQETLKYQDAFYGERSTQODIYAGSVQPIRLHLEGGN	120
Qy	121	AS-VLAYGPTGAGKTHTMLGSPQPGVTPRALMDLLQLTREBAGRFPWALSVTMSYLEI	179
Db	121	AKWVLYGPTGAGKT-THAGQPRATGWDPAGSHGPPAAHKGGCGRGPAMGLSVTMSYLEI	179
Qy	180	YOEKVLDDLDPASGDLVIREDCRGNILIPGLSOKPISFPADPERHFLPASNRRTVGATRL	239
Db	180	YOEKVLDDLDPASGDLVIREDCRGNILIPGLSOKPISFPADPERHFLPASNRRTVGATRL	239
Qy	240	NQRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDNRRRTGNKGLRKESGAIN	299
Db	240	NQRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGESEDNRRRTGNKGLRKESGAIN	299
Qy	300	SLFVLGKVVDAALNQGPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERRFYLDVTVAL	359
Db	300	SLFALGKVVDAALNQGPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERRFYLDVTVAL	359
Qy	360	NFAARSKEVIN 370	
Db	360	NFAARSKEVIN 370	



Db 150 SRILVFRLRPMGKERENGSCCVKVLNKRVDVLTFTNENDYLRKLRLVRHFTDSS 209  
Qy 94 YGERSTQDIYAGSQPILRHLLLEGONASVLAAYGPTGAKTHMLGSPQPGVIPRALMD 153  
Db 210 PFTTIOQEVSTTTGDIIVLEGRNSVFCYGATGAGKTYMLGTWENPGVWVLAIKD 269  
Qy 154 LQLTREBAGRPPWALSVTMSYLEIYQEKVLDDLPASGDIVIREDCRGNILIPGLSQK 213  
Db 270 LPAKVRQSLDNH---VHLSYLEVYNETVRDLSLSPGR-PLILREDKQ----VWALLQR 321  
Qy 214 PISSPADFERHFLPASRRTVTGATRLNQRSSSHAVLLVKVDQRLRAPPFR--OREGKLY 271  
Db 322 -----GNQNRTEPTRCNETSRSHAILQVIVIEYKTRDASMNIIISRVGKLS 367  
Qy 272 LIDLAGEEDNRRTGNKGLRKESGAINTSFLVLGVKVDALNOGLPRVPRDSKLTLLQD 331  
Db 368 LIDLAGEERALADQRTURLEGANINRSLALSSCINALVEGKHIPYRNSKLTQLLKD 427  
Qy 332 SLGSGSAHSILIANIAPERRFYLDVTVSALNFAARSKEV 368  
Db 428 SLGSGCNTVMIANISPSQSFGTQNTLHWADRAKEI 464

## RESULT 5

T18277  
kinesin heavy chain - slime mold (Dictyostellium discoideum)  
C:Species: Dictyostellium discoideum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18277  
R:deHostos, E.L.; McCaffrey, G.; Vale, R.D.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: Z18853  
A:Accession: T18277  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1254 <DEH>  
A:Cross-references: UNIPROT:Q94463; EMBL:U41289; NID:g1526990; PID:g1526991; PIDN:AAB077

Query Match 29.1%; Score 547; DB 2; Length 1254;  
Best Local Similarity 36.3%; Pred. No. 2.2e-35;  
Matches 122; Conservative 76; Mismatches 126; Indels 12; Gaps 5;  
Qy 39 PPPARVRVRLRPVDPDTAGASDPPCVRGMDSCSLEI-ANWRNHQETLKYQDAFYGER 97  
Db 24 PVSSNIRVVCVRPLTELEKGRNEHSIVHFPDSKSIIRANGP-----QTFDFRIFYQ 77  
Qy 98 STQDDIYAGSQPILRHLLLEGONASVLAAYGPTGAKTHMLGSPQPGVIPRALMDL--- 154  
Db 78 ETQSQIFEDVAEPIVNDPLDGVHGTIIAYGQTASGKTFMTVGDPSHGIIIPRVIESIFVG 137  
Qy 155 LQLTREBAGRPPWALSVTMSYLEIYQEKVLDDLPASGDIVIREDCRGNILIPGLSQK 214  
Db 138 ISKMRKDT-SLSLAFCLKISALELYNEKLYDYTASKSNIRHKGQIYVEGISIV 196  
Qy 215 ISSPADFERHFLPASRRTVTGATRLNQRSSSHAVLLVKVDQRLRAPPFRQEGKLYLD 274  
Db 197 ITSIEEAYFLNISNNRAIATKMSAASSHSHVMIELSQ-QNLSMESSKISKLFVLD 255  
Qy 275 LAGSEDNRRNTGNKGLRKESGAINTSFLVLGVKVDALNOGLPRVPRDSKLTLLQDLSLG 334  
Db 256 LAGSDISHTKGAEGDMQEMQEAKNINLSLALGKVINALTTCGVNYPYRDSKLTVLQDLSLG 315  
Qy 335 GSAHSILIANIAPERRFYLDVTVSALNFAARSKEVIN 370  
Db 316 GNSKTSLIINGSPNNNEHEHTITTLQFGTRAKTIKN 351

## RESULT 6

T47525  
kinesin-related protein-like - Arabidopsis thaliana  
N:Alternate names: protein F1612.60  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47525  
R:Jordan, N.; Banerdt, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.;  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24468  
A:Accession: T47525  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1058 <JOR>  
A:Cross-references: UNIPROT:Q9LZU5; EMBL:AL162459  
A:Experimental source: cultivar Columbia; BAC clone F1612  
C:Genetics:  
A:Map position: 3  
A:Introns: 56/2; 99/3; 151/3; 310/3; 343/2; 395/3; 419/1; 444/3; 462/3; 489/3; 527/1; 55/5  
A:Note: F1612.60  
C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology

Query Match 28.6%; Score 536.5; DB 2; Length 1058;  
Best Local Similarity 33.1%; Pred. No. 1.2e-34;  
Matches 134; Conservative 77; Mismatches 139; Indels 55; Gaps 9;  
Qy 6 STQQRREMAAASAA-----AISGAGRCRLSKIGATRRPPPAR-----VRVAVRLRPVDGT 57  
Db 3 SIQQRGGIVLSLPAQTTPRSSDKSARESRSESNSNTRNDKEKGVNVQVILRCRFLSEDE 62  
Qy 58 AGASDPPCVRGMDSCSLEIANWRNHQETLKYQ-----FDFVGERSTOODIVA 105  
Db 63 ARIHTPVVI-----SCN-----ENRREVAQTSTAGKHIDRHFAFDKVFGPASQOKDLVD 112  
Qy 106 GSVQPIRLHLLLEGONASVLAAYGPTGAKTHMLGAS-----PQPGVIPRALMDLQL 157  
Db 113 QAICPIVFEVLEGYNCTIFAYGQTGKTYMEGARKNGEFPDSAGVIPRAVKQIPDI 172  
Qy 158 TREBAGRPPWALSVTMSYLEIYQEKVLDDLP-----ASGDLVIREDCRGNILI 207  
Db 173 LEAQGAEE-----YSMKVTFLEYNEISDILLAPETIKFVDEKSKSIALMEDKGSVFV 227  
Qy 208 PGLSQKPTSSPADFERHFLPASRRTVTGATRLNQRSSSHAVLLVKVDQRLRAPPFRQ-- 265  
Db 228 RGLSEEVSTANEIYKILEKGSARTRTAETLUNKQSSHSIFSITTHIKNTPEGEEMI 287  
Qy 266 REGKLYLDLAGSEDNRRNTGNKGLRKESGAINTSFLVLGVKVDALNOGLPRVPRYRDSKL 325  
Db 288 KCGKLNLDVLAGSENISRSAGAREGAREAGEINKSLTLGRVINALVSHSGHIPYRDSKL 347  
Qy 326 TRLLQDSLGSSAHSILIANIAPERRFYLDVTVSALNFAARSKEVIN 370  
Db 348 TRLRLSGLGKTKCTVIATISPSIHCLTEETLSTLDYAHRAKNIKN 392

## RESULT 7

T40594  
probable kinesin-like protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T40594; T39531  
R:Purnelle, B.; Goffeau, A.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, May 1998  
A:Reference number: Z21939  
A:Accession: T40594  
A:Molecule type: DNA  
A:Residues: 1-784 <PUR>  
A:Cross-references: UNIPROT:O59751; EMBL:AL023587; PIDN:CAA19043.1; GSPDB:GN00067; SPDB:  
A:Experimental source: strain 972h-; cosmid c649  
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.  
submitted to the EMBL Data Library, July 1998  
A:Reference number: Z21861  
A:Accession: T39531  
A:Molecule type: DNA  
A:Residues: 464-784 <WOO>  
A:Cross-references: EMBL:AL031154; PIDN:CAA20063.1; GSPDB:GN00067; SPDB:SPBC1685.15C  
A:Experimental source: strain 972h-; cosmid c1685  
C:Genetics:





[illegible]

346 PSLGLGKALPRVRSKLLILVDLAGSERINKSTGDGHMIEEARFNLNLSLTSLGKCINNALAE 405

QY    314 GLRPVPYRDSKLTRLQLDSSLGGSANHSILIANTIAPEPPFYLDVTVSALNFAARSEKEVIN 370  
       :                        |                    ||                    |                    |                    |                    |                    |                    |                    |                    |  
Db    406 GSSHIPTRDSKLTRLRLDSFGGSARTSLIITIGPSARVHAETTSTIMFCQRAMKIYN 462

RESULT 12

H84777

probable kinesin-related cytokeletonis protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 Htext\_change 09-Jul-2004

C;Accession: H84777

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
euss, D.; Moo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: AB84420; PMID:20083487; PMID:10617197

Query Match	27.9%;	Score	524;	DB	2;	Length	1056;
Best Local Similarity	36.4%;	Pred. No.	1.2e-33;	Mismatches	67;	Indels	44; Gaps
Matches	131;	Conservative	67;	Mismatches	118;	Indels	44; Gaps
Qy	44	VRVAVRLRPFDVGTAGADPPCVRGMDSCSLBIANWRN----	HQETLKYQPDAFYGERST	99			
Dd	13	VQVLIRCRPFSDEI--RSNAPQVLTCLDLQREVAVSQNIAKHIDRV--FTFDKVEGPSAQ	70				
Qy	100	QQDIYAGSVQPIRLHLLBEGONASVIAYGPTGAKGTHTMLGS-----	PPQGPGY	147			
Dd	71	QXLDLYDQAVPIVNVELEGFNCTIFAYGTGTGKTYYTMEGECCRSKSAPCGSLPAEAGVI	130				
Qy	148	PRALMDLIQLTREBGAQRPMWLSVTMSYLEIYQEKVLDLLDP-----	ASGDIV	196			
Dd	131	PRAVRKQIF-----DTLEGGQAAYSVKVFLEYLNBEETDLLAPEDLSRAABEEKOKPKUP	185				

QY 197 IREDRCGNILPGLSOKPISS-----FADPERHFLPASNRKTVGATRLNQRSSRSHAVLLV 252  
Db 186 LMEDEGKGVLRGLEEIVSANEITFLER-----GSSKRRTAEFLNKQSSRSHLSFSI 241  
QY 253 KVDQERLAPFRQ--REGKLYLIDLAGEEDNRRTGNKGLRLKESGAINTSFLVLGKVDDA 310  
Db 242 THIKKATPEGBELIKGKLNVLVLAGSENISRSGARDGRAREAGEINKSLTLTLGRVISA 301  
QY 311 LNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 370  
Db 302 LVEHLGHVPYRDSKLTLLRDSLSGRTKTCIIATVSPAVHCLTEETLTDYAHRAKNKN 361

## RESULT 13

T02017  
kinesin-related protein TKRP125 - common tobacco  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T02017  
J:Cell Sci. 110, 179-189, 1997  
R:Asada, T.; Kuriyama, R.; Shibaoaka, H.  
A:Title: TKRP125, a kinesin-related protein involved in the centrosome-independent organ  
A:Reference number: Z14490; MUID:97196959; PMID:9044048  
A:Accession: T02017  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1006 <ASA>  
A:Cross-references: UNIPROT:O23826; EMBL:D83711; NID:g2582970; PIDN:BAA23159.1; PID:g258  
C:Genetics:  
A:Note: TKRP125  
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology  
F:10-361/Domain: kinesin motor domain homology <KNOT>

Query Match 27.8%; Score 521.5; DB 2; Length 1006;  
Best Local Similarity 35.6%; Pred. No. 1.8e-33;  
Matches 128; Conservative 69; Mismatches 118; Indels 45; Gaps 9;

QY 44 VRVAVRLPFPVDTAGSDPPCVGMD-----SCSLIANWRNHQETLYKYQFDFAFYGERS 98  
Db 10 VQVLLRCRFPFNDLRNAPQVTCNDYQREAVVSQNIAG--KHIDRI-FTFDKVFQPSA 66  
QY 99 TQODIYAGSVOPILRHLLGEGNASVLAYGPTGAGKTHMLGS-----PEQPGV 146  
Db 67 QORDLYDQAIIVPNEVELEGNCITFAYGQGTGKTYMEGECKRSKSGPNEGELPQEA 126  
QY 147 IPRALMDLLQLTRREGAERPWALSVTMSYLEIYQEKVLDLLDP-----ASGLV 196  
Db 127 IPRAVKQVDTLESQNAE-----YSVKVTFLEYNEETDILLAPEDLKVALEDROKKQLP 181  
QY 197 TREDRCGNILPGLSOKPISS-----FADPERHFLPASNRKTVGATRLNQRSSRSHAVLLV 252  
Db 182 LMEDEGKGVLRGLEEIVTANEITFLER-----GSSKRRTAEFLNKQSSRSHLSFSI 237  
QY 253 KVDQERLAPFRQ--REGKLYLIDLAGEEDNRRTGNKGLRLKESGAINTSFLVLGKVDDA 310  
Db 238 THIKKATPEGBELIKGKLNVLVLAGSENISRSGARDGRAREAGEINKSLTLTLGRVINA 297  
QY 311 LNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 370  
Db 298 LVEHLGHVPYRDSKLTLLRDSLSGRTKTCIIATVSPAVHCLTEETLTDYAHRAKNKN 357

## RESULT 14

B84687  
probable kinesin-like spindle protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: B84687  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84687  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1076 <STO>  
A:Cross-references: UNIPROT:Q9S1B3; GB:A8002093; NID:g4580395; PIDN:AAD24373.1; GSPDB:GN;  
C:Genetics:  
A:Gene: At2g28620  
A:Map position: 2  
C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology

Query Match 27.6%; Score 517.5; DB 2; Length 1076;

Best Local Similarity 32.6%; Pred. No. 4.1e-33;  
Matches 126; Conservative 74; Mismatches 152; Indels 35; Gaps 7;

QY 6 STQQRREMAAASAAISGACRLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPC 65  
Db 21 STEKSNRDFRVS-----NSNSNPVSKNEKEKGVNIQIVVRCRPF-NSEETRLQTPA 71  
QY 66 VRGMDSCSLEIA---NWRNHQETLYKYQFDFAVGERSTOODIYAGSVOPILRHLLGQNAS 122  
Db 72 VLTCDNRKKEVAVAGNIAGKQIDKTFLEKVFQPTSQOKDLYHQAVSPIVEFVLDGYNCT 131  
QY 123 VLAYGPTGAGKTHMLGS-----PEQPGVIPRALMDLLQLTRREGAERPWALSVTM 174  
Db 132 IPAYGQGTGKTYMEGARKNGEIPSDAGVIPRAVKQIFDILEAQSA---EYSLKV 187  
QY 175 SVLEYIYQEKVLDLL-----DPASGLVIREDCRGNILPGLSOKPISSPADFERHF 225  
Db 188 SFLEYNEELTDLLAPEETKFADDKSKKPLALMEDGKGVFVRGLEEIVTADIEYKVL 247  
QY 226 LPASNRKTVGATRLNQRSSRSHAVLLVKVDORERLAPFRQ--REGKLYLIDLAGEEDNR 283  
Db 248 EKGSAKRRTAEFLNKQSSRSHSIFSVTHIKECTPEGEIEVKSGKLNVLVLAGSENISR 307  
QY 284 TGNKGLRLKESGAINTSFLVLGKVVDALNQGFLRPVYRDSKLTLLQDSLGSAHSILIA 343  
Db 308 SGAREGRAREAGEINKSLTLGRVINALVEHSGHIPYRESKLTLLRDSLSGRTKTCVIA 367  
QY 344 NTAPERFVLDTVSALNFAARSKVIN 370  
Db 368 TVSPSVHCLTEETLTDYAHRAKNKN 394

## RESULT 15

S38982  
kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)  
N:Alternate names: kinesin-2 chain A; KEP (85/95) 85k chain  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S38982; S72551  
R:Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.  
Nature 366, 268-270, 1993  
A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.  
A:Reference number: S38982; MUID:94050179; PMID:8232586  
A:Accession: S38982  
A:Molecule type: mRNA  
A:Residues: 1-699 <COL1>  
A:Cross-references: UNIPROT:P46872; EMBL:Li6993; NID:g295245; PIDN:AAA16098.1; PID:g2952  
A:Accession: S72551  
A:Molecule type: protein  
A:Residues: 2-5, 'X', 7-11, 59-64, 125-132, 222-226, 'X', 228-230 <COL2>  
C:Complex: heterotrimer of a 115k chain and two kinesin-related chains of 95k (PIR:S586)  
C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
C:Keywords: ATP; heterotrimer; microtubule binding; nucleotide binding; P-loop  
F:11-348/Domain: kinesin motor domain homology <KNOT>  
F:97-104/Region: nucleotide-binding motif A (P-loop)  
F:103/Binding site: ATP (Lys) #status predicted

Query Match 27.4%; Score 514.5; DB 1; Length 699;

Best Local Similarity 37.6%; Pred. No. 3.9e-33;  
Matches 127; Conservative 61; Mismatches 135; Indels 15; Gaps 8;



This Page Blank (uspto)



```

Qy 314 GLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARKEV 368
Db 304 ---HIPYRNSKLTLLQDSLGSGCQTIIMIAAVSPSSVFYDDTYNTLKYANRAKDI 355

RESULT 2
US-10-952-698-164
; Sequence 164, Application US/10952698
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF BLADDER CANCER. COMPOSITIONS AND METHODS
; TITLE OF INVENTION: OF SCREENING FOR MODULATORS OF BLADDER CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/952,698
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 164
; LENGTH: 1380
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-952-698-164

Query Match 28.0%; Score 526; DB 6; Length 1380;
Best Local Similarity 37.2%; Pred. No. 6.3e-41;
Matches 128; Conservative 66; Mismatches 128; Indels 22; Gaps 7;

Qy 44 VRVAVRLRPFDV--GTAGASDPPCVRGMDSCSLEIANRNHQTLLKYQFDAYGERSTQ 101
Db 27 IKVFRIRPPAERSGSADGEQNLCLSVLSLSLRL---HSNPEKPTFFDHVADVDITQE 83

Qy 102 DIYAGSVQPIRLHLLLEGONASVLAYGPTGAKTHMLGSPQEP-----GVIPRA---L 151
Db 84 SVFATVAKSIVSFCMSGYNGTIFAYGQTGSGKTFMTMGPSSEDNFNSHNLGRVIPSPEYL 143

Qy 152 MDLLQLFREEGACGRPWALSVTMSYLEIYQSKVLDLDPASGDVIVREDCKGNILIPGLS 211
Db 144 FSLIDREKEKAGAGKSPFCK--CSFIEIYNEQIYDLDSAGLYLREHIKGVFVVGAV 201

Qy 212 QKPISSAFADPRHFLPASRNRTVGATRLNQRSSSHAVLVKVDORERLAPP--RQREGKL 270
Db 202 EQVTSAAEAQVLSGGRNRRVASTSNRRESSSHAVFTTISMEKSNEIVNIRTSLL 261

Qy 271 YLIDLAGSEDRNRTGNKGLRKESGAINTSFLVGLKVVDAL----NOGLPRVPYRDSKLT 326
Db 262 NLVDLAGSERQKQTHAEGMRLEKAGNINRSLSCLGVITALVDVGNKGQRHVCYRDSKLT 321

Qy 327 RLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARKEVIN 370
Db 322 FLRLDSLGNKARTAIANVHFGSCFCGETLTLNPAQAKLIKXN 365

RESULT 3
US-10-948-973-2
; Sequence 2, Application US/10948973
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1043
; CURRENT APPLICATION NUMBER: US/10/948,973
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: US/10/126,205
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/723,595
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 361

```

```

; TYPE: PRT
; ORGANISM: Human
US-10-948-973-2

Query Match 22.5%; Score 423; DB 6; Length 361;
Best Local Similarity 32.8%; Pred. No. 4e-32;
Matches 121; Conservative 59; Mismatches 131; Indels 58; Gaps 11;

Qy 44 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLEIANRNHQTLLK-- 87
Db 8 IRVFCRVRLPGEPTPPGGLLFPSPGPGSDPP-----TRLSRSDERRGTLGA 60

Qy 88 -----YQFDAYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAKTHML- 138
Db 61 PAPTRHDFSFDRVPFPGSGQDEVEF-EIAMLVQSALDGYPCIFAYGQTGSGKFTMEG 119

Qy 139 ---GSPEQPGVTPRALMDLLQLTREBAGRPWALSVTMSYLEIYQSKVLDLDPAS--- 192
Db 120 GPGGDPQLEGLIPRALRHLFSVAQELSGQG--WTYSFVASYVEIYNETVRDLLATGTRKG 177

Qy 193 ---GDLVIREDCRGN--ILIPGLSQPISSFADFERHFLPASRNRTVGATRLNQRSSSHA 248
Db 178 QGGECEIRRAGPGSEELVTNARYVPVSCKEVDALLHLARONRAVARTAQNERSRSHS 237

Qy 249 VLLVKVDQERLAPFQREGKLYLIDLAGSEDRNRTGNKGL-----RLKESGAINTS 300
Db 238 VFQLQI--SGEHSRGLQCGAPLSLVDLAGSE---RLDPGLALGPGERERLRETQAINSS 292

Qy 301 LFLVGLKVVDALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDITVSALN 360
Db 293 LSTGLGLVIMALSNKESHVYRNSKLTLYLLQNSLGGSAKMLMFVNISPLEENVSESLNLR 352

Qy 361 FAARKEVIN 369
Db 353 FASKVNQCV 361

RESULT 4
US-10-948-973-6
; Sequence 6, Application US/10948973
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1043
; CURRENT APPLICATION NUMBER: US/10/948,973
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: US/10/126,205
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/723,595
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Human
US-10-948-973-6

Query Match 22.5%; Score 423; DB 6; Length 369;
Best Local Similarity 32.8%; Pred. No. 4.1e-32;
Matches 121; Conservative 59; Mismatches 131; Indels 58; Gaps 11;

Qy 44 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLEIANRNHQTLLK-- 87
Db 7 IRVFCRVRLPGEPTPPGGLLFPSPGPGSDPP-----TRLSRSDERRGTLGA 59

Qy 88 -----YQFDAYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAKTHML- 138
Db 60 PAPTRHDFSFDRVPFPGSGQDEVEF-EIAMLVQSALDGYPCIFAYGQTGSGKFTMEG 118

```

```

QY 139 ---GSPEQGVIPRALMDLLQLTREGGAEGRPWALSVTWSYLEIYOEKVLDLLDPAS---- 192
Db 119 GPGGDPQLLEGLIPRALRHLFSVAQELSGQG--WTYSFVASYVEIYNETVRDLATCTRRG 176
QY 193 --GDLVIREDCRGN--ILIPGLSQKPIGSFADFERRHFLPASNRRTVGATRLNQRSSRSHA 248
Db 177 QGGECEIRRAGPGSEELVTWARYVPVSCKEVDALLHLARQNRVAVTAQNERSRSHS 236
QY 249 VLLVKVDORERLAPRQREGKLYLIDLAGESENRTGNKGL-----RLKESGAINTS 300
Db 237 VFQLOQI-SGEHSSRGLOCGAPLSLVDLAGE----RLDPGLALGPGERERLRETQAINSS 291
QY 301 LFLVLGKVVDALNOGLPRVPYPRDSKLTLLQDSLGGSAHSILIANIAPERRFVLDTVSALN 360
Db 292 LSTGLGLVIMALSINKESHVPYRNSKLYTLLONSLGGSAKMLMPVNISPLEENVSELSNLR 351
QY 361 PAARSKEVI 369
Db 352 FASKVNQC 360

RESULT 5
US-10-948-973-4
; Sequence 4, Application US/10948973
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1043
; CURRENT APPLICATION NUMBER: US/10/948, 973
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: US/10/126, 205
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/723, 595
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/295, 612
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-10-948-973-4

```

```

Qy      361 FAARSEKVI 369
      ||:: :
Db      353 FASKVQCV 361

RESULT 6
US-10-951-582-4
; Sequence 4, Application US/10951582
; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/595,424
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Human
US-10-951-582-4

```

```

; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/595,424
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Human
; US-10-951-582-6

```

```

Query Match      21.0%; Score 394; DB 6; Length 658;
Best Local Similarity 27.3%; Pred. No. 5e-29;
Matches 117; Conservative 64; Mismatches 127; Indels 120; Gaps 13;

QY 44 VRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIAN-----WR-----NHQETLK 87
DB 24 VGVYCRVRPL-----GFPD-----QECIEVINNTTVQLHTPEGYRLNRNGDYKET-Q 70

QY 88 YQFDIFYGERSTOODIYAGSVQPIRLHLEQONASVLAIGPTGAGKTHTMLGSPQPGVI 147
DB 71 YSFQKQVGTHTTQKELFDVVANPLVNDLHGKNGLLFTYGVGTSGKTHTMLGSPGEGLL 130

QY 148 PRAL-----MD-----LLQLTREGAGRPWALS-----I 197
DB 131 PRCLDMIFNSIGSFQAKRYVFKSNDNRNSMDIQCEVDALLERQKREAMPNPKTSSSKRQVD 190

QY 172 -----VTMSYLEIQEKVLDLLDPASGDLV-----I 197
DB 191 PEFADMITVQEFKAEVEDESDSVGVFVSIEIYNNYIYDLLEVPDPKPKPPQSKLL 250

QY 198 REDCRGNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQR 257
DB 251 REDKNHNMVAGTEVEVKSTEEAFVFWRGQKKRIANTHLNRESRSHSVFNKILVQ- 309

QY 258 ERLAP-----FQRE-----GKLYLIDLAGEDNRRTGNKGLRLKESGAINTSFLVLG 305
DB 310 ---APLDADGDNVLQEKQITISQLSLVDLAGSERTNRTAEGNRLREAGNIQSLMTLR 366

QY 306 KVVDAALNOGL-----PRVYRDSKLTRELQDSLGSAHSILIANIAPERFYLDITVSALN 360
DB 367 TCMVDVLRNQMVGTKNMPYRDSKLTFLFKNYPDGEKGVRMIVCVNPKAEDYEENLQVMR 426

QY 361 FAARSKEV 368
DB 427 FAEVTOEV 434

```

```

RESULT 8
US-10-951-582-2
; Sequence 2, Application US/10951582
; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582

```

```

; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/595,424
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Human
; US-10-951-582-2

Query Match      20.8%; Score 390; DB 6; Length 433;
Best Local Similarity 27.2%; Pred. No. 6.5e-29;
Matches 116; Conservative 64; Mismatches 127; Indels 120; Gaps 13;

QY 44 VRVAVRLRPFDGTAGASDPPCVRGMDSCSLEIAN-----WR-----NHQETLK 87
DB 24 VGVYCRVRPL-----GFPD-----QECIEVINNTTVQLHTPEGYRLNRNGDYKET-Q 70

QY 88 YQFDIFYGERSTOODIYAGSVQPIRLHLEQONASVLAIGPTGAGKTHTMLGSPQPGVI 147
DB 71 YSFQKQVGTHTTQKELFDVVANPLVNDLHGKNGLLFTYGVGTSGKTHTMLGSPGEGLL 130

QY 148 PRAL-----MD-----LLQLTREGAGRPWALS-----I 197
DB 131 PRCLDMIFNSIGSFQAKRYVFKSNDNRNSMDIQCEVDALLERQKREAMPNPKTSSSKRQVD 190

QY 172 -----VTMSYLEIQEKVLDLLDPASGDLV-----I 197
DB 191 PEFADMITVQEFKAEVEDESDSVGVFVSIEIYNNYIYDLLEVPDPKPKPPQSKLL 250

QY 198 REDCRGNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQR 257
DB 251 REDKNHNMVAGTEVEVKSTEEAFVFWRGQKKRIANTHLNRESRSHSVFNKILVQ- 309

QY 258 ERLAP-----FQRE-----GKLYLIDLAGEDNRRTGNKGLRLKESGAINTSFLVLG 305
DB 310 ---APLDADGDNVLQEKQITISQLSLVDLAGSERTNRTAEGNRLREAGNIQSLMTLR 366

QY 306 KVVDAALNOGL-----PRVYRDSKLTRELQDSLGSAHSILIANIAPERFYLDITVSALN 360
DB 367 TCMVDVLRNQMVGTKNMPYRDSKLTFLFKNYPDGEKGVRMIVCVNPKAEDYEENLQVMR 426

QY 361 FAARSKE 367
DB 427 FAEVTOE 433

RESULT 9
US-10-951-582-8
; Sequence 8, Application US/10951582
; GENERAL INFORMATION:
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1041
; CURRENT APPLICATION NUMBER: US/10/951,582
; CURRENT FILING DATE: 2004-09-27
; PRIOR APPLICATION NUMBER: US/10/143,563
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US/09/723,630
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/595,424
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/314,464

```



```
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Human
US-10-951-582-8

Query Match
Best Local Similarity 19.8%; Score 371; DB 6; Length 960;
Matches 110; Conservative 63; Mismatches 118; Indels 110; Gaps 13;

QY 72 CSLEIAN-----WR-----NHQETLYQFDAYGERSTQDDIYAGSVQPIRLHL 115
Db 40 CCIEVINNTTVQLHTPEGYRLNRNGDYKET-QYSFKQVFGTHTTQKELFDVVANPLVNDL 98
QY 116 LEGQNASVLAYPEGTAGKTHMLGSPQGVIPRAL----- 151
Db 99 IHKRGKLLFTYGVTSKTHMTGSPGEGLLPRCLDMIFNSIGSFQAKRYVFKSNDNRNS 158
QY 152 MD-----LLQLTREGAGRPWALS-----VTM 174
Db 159 MDIQEVDALLERQKREAMPNPKTSKQKQVDPEDPMITVOEFCKAEVDEDSVYGVPV 218
QY 175 SYLEIYQEKVLDL-----DPASGDLVIREDC-----RGNILIPGLSQKPISSPADFERH 224
Db 219 SYIEIYNNVIYDLLEVPDPINPL-HNLNCFVKIKHNMVYAGTEVEVKSTEEAFV 277
QY 225 FLPASRNTVGNTRLNQSSRSRSHAVLLVKVQDQERLAP-----FRQRE-----GKLYL 272
Db 278 FWRGQKRRRIANTHLNRSRSHSVFNKLQV-----APLDADGDNVLOEKEQITISQLSL 333
QY 273 IDLAGSDNRRTGKGLRLKESGANTSLFVLGKVVDAALNQL-----PRVPEYRDSKLT 327
Db 334 VDLASERTNTRAGNRLUREAGNIQSLMTRTCDVLRENMQYGTNNQVYRDSKLT 393
QY 328 LLQDSLGSAHSILIANIAPERFRYLDVTSALNFAARKEV 368
Db 394 LFNKYFDGEGKVMIVCNPKAEDYEENLQVRAEVTQEV 434

RESULT 10
US-10-220-366A-19691
; Sequence 19691, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 19691
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-19691

Query Match
Best Local Similarity 11.0%; Score 206.5; DB 6; Length 137;
Matches 55; Conservative 15; Mismatches 44; Indels 5; Gaps 2;

QY 244 SRSHAVLLVKVDQERLAPFRQ--REGKLYLDLAGSDNRRTGKGLRLKESGANTSL 301
Db 16 SGTHAVFOFFLKKQGRVPLGTQAVQAKMSLIDLAGSRASSTHAKGERLREGANINSL 75
QY 302 FVLGKVVDAALNQLPR---VPRYRDSKLTLLQDSLGSAHSILIANIAPERFRYLDVTS 357
Db 16 SGTHAVFOFFLKKQGRVPLGTQAVQAKMSLIDLAGSRASSTHAKGERLREGANINSL 75

; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Human
US-10-951-582-8

Query Match
Best Local Similarity 19.8%; Score 371; DB 6; Length 960;
Matches 110; Conservative 63; Mismatches 118; Indels 110; Gaps 13;

QY 72 CSLEIAN-----WR-----NHQETLYQFDAYGERSTQDDIYAGSVQPIRLHL 115
Db 40 CCIEVINNTTVQLHTPEGYRLNRNGDYKET-QYSFKQVFGTHTTQKELFDVVANPLVNDL 98
QY 116 LEGQNASVLAYPEGTAGKTHMLGSPQGVIPRAL----- 151
Db 99 IHKRGKLLFTYGVTSKTHMTGSPGEGLLPRCLDMIFNSIGSFQAKRYVFKSNDNRNS 158
QY 152 MD-----LLQLTREGAGRPWALS-----VTM 174
Db 159 MDIQEVDALLERQKREAMPNPKTSKQKQVDPEDPMITVOEFCKAEVDEDSVYGVPV 218
QY 175 SYLEIYQEKVLDL-----DPASGDLVIREDC-----RGNILIPGLSQKPISSPADFERH 224
Db 219 SYIEIYNNVIYDLLEVPDPINPL-HNLNCFVKIKHNMVYAGTEVEVKSTEEAFV 277
QY 225 FLPASRNTVGNTRLNQSSRSRSHAVLLVKVQDQERLAP-----FRQRE-----GKLYL 272
Db 278 FWRGQKRRRIANTHLNRSRSHSVFNKLQV-----APLDADGDNVLOEKEQITISQLSL 333
QY 273 IDLAGSDNRRTGKGLRLKESGANTSLFVLGKVVDAALNQL-----PRVPEYRDSKLT 327
Db 334 VDLASERTNTRAGNRLUREAGNIQSLMTRTCDVLRENMQYGTNNQVYRDSKLT 393
QY 328 LLQDSLGSAHSILIANIAPERFRYLDVTSALNFAARKEV 368
Db 394 LFNKYFDGEGKVMIVCNPKAEDYEENLQVRAEVTQEV 434

RESULT 11
US-10-220-366A-14894
; Sequence 14894, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 14894
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-14894

Query Match
Best Local Similarity 9.7%; Score 183; DB 6; Length 134;
Matches 47; Conservative 19; Mismatches 49; Indels 2; Gaps 1;

QY 172 VTMSYLEIYQEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRN 231
Db 18 VMSDYLEIDSEMIRODLNTFPVHLELQDDSSGVIQAGITEASTNAKEIMHLLMKGNQ 77
QY 232 RTVGATRLNQRSSRSRSHAVLLVKVQDQERLAPFRQ--REGKLYLDLAGSDNRRTGN 286
Db 78 RTQEPNTAANTSSRSRSHAVLLQVTVQRSRVKNLQEVROGRFLPMIDLAGSERASQTQN 134

RESULT 12
US-10-220-366A-26476
; Sequence 26476, Application US/10220366A
; GENERAL INFORMATION:
; APPLICANT: HYSEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 26476
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-26476

Query Match
Best Local Similarity 9.6%; Score 180; DB 6; Length 134;
Matches 42; Conservative 21; Mismatches 53; Indels 10; Gaps 2;

QY 41 PAR-----VRVAVRLRPFPVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLYK-----QF 90
Db 1 PAREMEKFRVCVRKPLGMRVRRGEINIITVEDKETLLVHEKKAVDLTQVILQHVFF 60
QY 91 DAFYGERSTQDDIYAGSVQPIRLHLLEQNASVLAYPEGTAGKTHMLGSPQGVIPRA 150
Db 61 DEVFGEACTNQDVYMKTTPLIQHIFNGGNATCFAYGQTGAGKTYTMTIGTHENPGLYALA 120
```

QY 151 LMDLLQ 156  
 Db 121 AKUIFR 126

RESULT 13  
 US-10-220-366A-15821  
 ; Sequence 15821, Application US/10220366A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HYSEQ, INC  
 ; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
 ; FILE REFERENCE: 21272-042  
 ; CURRENT APPLICATION NUMBER: US/10/220,366A  
 ; CURRENT FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: 09/577,409  
 ; PRIOR FILING DATE: 2000-05-18  
 ; PRIOR APPLICATION NUMBER: 09/515,126  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 27802  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 15821  
 ; LENGTH: 139  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(139)  
 ; OTHER INFORMATION: Xaa = any amino acid or nothing

US-10-220-366A-15821  
 Query Match 9.3%; Score 174; DB 6; Length 139;  
 Best Local Similarity 43.2%; Pred. No. 2.1e-09;  
 Matches 41; Conservative 17; Mismatches 25; Indels 12; Gaps 2;

QY 288 GIRLKESAINTSFLVLKGVVDAL---NOGLPR-----VYRDSKLTLLIQDSLGG 335  
 Db 6 GVRLKEGGINTEFVTGNGVNSALADFSQDASRYSKKKQVLVYRDSVMTWLLKDSLGG 65

QY 336 SAHSILIANIAPERFRFYLDITVSALNFAARKEVIN 370  
 Db 66 NSKALMATISPADVNYGETLSTLYANRAKNIIN 100

RESULT 14  
 US-10-732-923-6841  
 ; Sequence 6841, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 6841  
 ; LENGTH: 911  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana

US-10-732-923-6841  
 Query Match 6.0%; Score 113; DB 6; Length 911;  
 Best Local Similarity 21.4%; Pred. No. 0.018;  
 Matches 83; Conservative 69; Mismatches 128; Indels 108; Gaps 20;

QY 2 AAGSTQORRREMAAASAAAGRCRLSKIGATRRPP---PA-----RVRVAVR 49  
 Db 52 SAGG-----ENAAQSAERVINQA---LKKLPSPQPPDDIPASSSLIKVIRRAQAQK 101

QY 50 LRP-----FVD-----GTAGA---SDPPCVRGMDSCSLEIANWRNHQ 83  
 Db 102 SRGDTHLAVDQIMGLLEDSQIRDLNVEGVATARKSEVEKLRGKGVESASGDTNF 161

QY 84 ETLKYQFDAYGERSTQODIYAGSVQPII-----RHLLEGQNASVLAYGPTGAGK 133  
 Db 162 QALK-----TYGRDLVEQ---AGKLDPVIGRDEIRRVVRLSRRYKNNPVLIGEPGVGK 213

QY 134 THTMLGSPQEQ--PGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEK---VLDLL 188  
 Db 214 TAVVEGLAQRIKVGDVNSLTDVRLISLDMG-----ALVAGAKYRGFEERLKSVLKEV 267

QY 189 DPASGDLVIREDCRGNILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSRSHA 248  
 Db 268 EDAGKVLIFIDEIHLVLGAGKTEGMDA-ANLFPKMLARGQLRCIGATTLEE--YRKV 324

QY 249 VLLVKVDQBERLAPFQREGKLYLD-----LAGSEDRRTTGNKGLRKESGAIN- 298  
 Db 325 -----EKDAAFERRFOQVVAEVPDITISILRLKE-KYEGHHGVRIQDRALINA 374

QY 299 ---TSLFVLG-----KVVDALNQGILPRV 318  
 Db 375 AQLSARYITGRHLDPKADLVDDEACANV 402

US-10-732-923-6842  
 ; Sequence 6842, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 6842  
 ; LENGTH: 911  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana

QY 151 LMDLLQ 156  
 Db 121 AKUIFR 126

RESULT 13  
 US-10-220-366A-15821  
 ; Sequence 15821, Application US/10220366A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HYSEQ, INC  
 ; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
 ; FILE REFERENCE: 21272-042  
 ; CURRENT APPLICATION NUMBER: US/10/220,366A  
 ; CURRENT FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: 09/577,409  
 ; PRIOR FILING DATE: 2000-05-18  
 ; PRIOR APPLICATION NUMBER: 09/515,126  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 27802  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 15821  
 ; LENGTH: 139  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(139)  
 ; OTHER INFORMATION: Xaa = any amino acid or nothing

US-10-220-366A-15821  
 Query Match 9.3%; Score 174; DB 6; Length 139;  
 Best Local Similarity 43.2%; Pred. No. 2.1e-09;  
 Matches 41; Conservative 17; Mismatches 25; Indels 12; Gaps 2;

QY 288 GIRLKESAINTSFLVLKGVVDAL---NOGLPR-----VYRDSKLTLLIQDSLGG 335  
 Db 6 GVRLKEGGINTEFVTGNGVNSALADFSQDASRYSKKKQVLVYRDSVMTWLLKDSLGG 65

QY 336 SAHSILIANIAPERFRFYLDITVSALNFAARKEVIN 370  
 Db 66 NSKALMATISPADVNYGETLSTLYANRAKNIIN 100

RESULT 14  
 US-10-732-923-6841  
 ; Sequence 6841, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 6841  
 ; LENGTH: 911  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana

US-10-732-923-6841  
 Query Match 6.0%; Score 113; DB 6; Length 911;  
 Best Local Similarity 21.4%; Pred. No. 0.018;  
 Matches 83; Conservative 69; Mismatches 128; Indels 108; Gaps 20;

QY 2 AAGSTQORRREMAAASAAAGRCRLSKIGATRRPP---PA-----RVRVAVR 49  
 Db 52 SAGG-----ENAAQSAERVINQA---LKKLPSPQPPDDIPASSSLIKVIRRAQAQK 101

QY 50 LRP-----FVD-----GTAGA---SDPPCVRGMDSCSLEIANWRNHQ 83  
 Db 102 SRGDTHLAVDQIMGLLEDSQIRDLNVEGVATARKSEVEKLRGKGVESASGDTNF 161

US-10-732-923-6842  
 ; Sequence 6842, Application US/10732923  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10/732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 6842  
 ; LENGTH: 911  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana

Search completed: November 5, 2004, 19:04:48  
Job time : 6.60933 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 62.7813 Seconds  
(without alignments)  
2081.726 Million cell updates/sec

Title: US-10-797-893-2

Perfect score: 1878

Sequence: 1 MAAGSTQQRREMAAASAA.....FYLDTSALNFAARSKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*

2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*

3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*

4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*

5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*

6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*

7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*

8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*

9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*

10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*

11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*

12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*

13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*

14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*

15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*

16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*

17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*

18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*

19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	370	16 US-10-797-893-2	Sequence 2, Appli
2	1878	100.0	490	15 US-10-334-143-80	Sequence 80, Appli
3	1873	99.7	512	16 US-10-797-893-4	Sequence 4, Appli
4	1763	93.9	346	16 US-10-797-893-6	Sequence 6, Appli
5	1763	93.9	460	9 US-09-925-300-1228	Sequence 1228, Ap
6	1763	93.9	487	16 US-10-797-893-8	Sequence 8, Appli
7	601	32.0	589	17 US-10-425-115-320209	Sequence 320209,
8	597.5	31.8	377	16 US-10-437-963-159031	Sequence 159031,
9	559	29.8	548	15 US-10-108-260A-2692	Sequence 2692, Ap
10	557.5	29.7	783	17 US-10-425-115-332919	Sequence 332919,
11	557.5	29.7	854	15 US-10-425-114-59708	Sequence 59708, A
12	544.5	29.0	776	16 US-10-437-963-103943	Sequence 103943,
13	528	28.1	383	9 US-09-883-096-5	Sequence 5, Appli

14	528	28.1	864	9 US-09-883-096-2	Sequence 2, Appli
15	526	28.0	375	15 US-10-332-089-4	Sequence 4, Appli
16	526	28.0	409	15 US-10-332-089-6	Sequence 6, Appli
17	526	28.0	774	14 US-10-369-493-2361	Sequence 2361, Ap
18	526	28.0	1388	14 US-10-146-473-82	Sequence 82, Appl
19	526	28.0	1388	15 US-10-173-999-32	Sequence 32, Appl
20	526	28.0	1388	15 US-10-332-089-2	Sequence 2, Appli
21	526	28.0	1388	15 US-10-188-832-164	Sequence 164, App
22	525.5	28.0	757	15 US-10-220-120-366	Sequence 366, App
23	525.5	28.0	757	15 US-10-363-829-316	Sequence 316, App
24	525	28.0	1030	15 US-10-425-114-62748	Sequence 62748, A
25	523.5	27.9	834	15 US-10-287-226-306	Sequence 306, App
26	521.5	27.8	1401	15 US-10-287-226-142	Sequence 142, App
27	520	27.7	1001	17 US-10-425-115-231231	Sequence 231231,
28	519.5	27.7	1826	16 US-10-408-765A-1477	Sequence 1477, Ap
29	518.5	27.6	1382	16 US-10-437-963-176714	Sequence 176714,
30	518	27.6	1115	15 US-10-287-226-302	Sequence 302, App
31	518	27.6	1232	14 US-10-116-712-670	Sequence 670, App
32	518	27.6	1232	16 US-10-408-765A-2153	Sequence 2153, Ap
33	518	27.6	1235	15 US-10-334-143-8	Sequence 8, Appli
34	517.5	27.6	1324	15 US-10-287-226-314	Sequence 314, App
35	516	27.5	992	17 US-10-739-930-7903	Sequence 7903, Ap
36	513.5	27.3	805	14 US-10-369-493-21903	Sequence 21903, A
37	513	27.3	1237	15 US-10-334-143-33	Sequence 33, Appl
38	510	27.2	672	16 US-10-408-765A-1664	Sequence 1664, Ap
39	509	27.1	1103	9 US-09-847-874A-1	Sequence 1, Appli
40	509	27.1	1103	14 US-10-458-162-1	Sequence 1, Appli
41	507	27.0	928	14 US-10-080-608A-23	Sequence 23, Appl
42	507	27.0	928	14 US-10-370-685-112	Sequence 112, App
43	507	27.0	1232	14 US-10-116-712-664	Sequence 664, App
44	507	27.0	1232	14 US-10-116-712-669	Sequence 669, App
45	504	26.8	338	9 US-09-883-096-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1

US-10-797-893-2

; Sequence 2, Application US/10797893

; Publication No. US20040142397A1

; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe

; TITLE OF INVENTION: Novel motor proteins and methods for

; FILE REFERENCE: their use

; FILE REFERENCE: 1044

; CURRENT APPLICATION NUMBER: US/10797,893

; CURRENT FILING DATE: 2004-03-09

; PRIOR APPLICATION NUMBER: US/09/724,224

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: 09/597,292

; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Human

; US-10-797-893-2

Query Match	100.0%	Score	1878;	DB	16;	Length	370;
Best Local Similarity	100.0%	Pred. No.	4.3e-177;				
Matches	370;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MAAGSTQQRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRP	FDGTAGA	60		
DB	1	MAAGSTQQRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVAVRLRP	FDGTAGA	60		
QY	61	SDPPCVRGMDSCSLEIANRNH	QETLYQFPAFYGERSTQODIYAGSVQPI	LRLHLEGGN	120		
DB	61	SDPPCVRGMDSCSLEIANRNH	QETLYQFPAFYGERSTQODIYAGSVQPI	LRLHLEGGN	120		
QY	121	ASVLAYGPTGAGTKHTMLGS	PEQCVIPRALMDLLQLTREEGAEGRPWALS	VTMSYLEIY	180		

Db 121 ASVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 180  
QY 181 QEKVLDLLDPASGLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLN 240  
Db 181 QEKVLDLLDPASGLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLN 240  
QY 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGEEDNRRTGNKGLRKESGAINTS 300  
Db 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGEEDNRRTGNKGLRKESGAINTS 300  
QY 301 LFVLGKVVVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 360  
Db 301 LFVLGKVVVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 360  
QY 361 FAARSKEVIN 370  
Db 361 FAARSKEVIN 370  
RESULT 2  
US-10-334-143-80  
; Sequence 80, Application US/10334143  
; Publication No. US20040009549A1  
; GENERAL INFORMATION:  
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1543  
; CURRENT APPLICATION NUMBER: US/10/334,143  
; PRIOR FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: 60/343,169  
; PRIOR FILING DATE: 2001-12-31  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-334-143-80  
Query Match 100.0%; Score 1878; DB 15; Length 490;  
Best Local Similarity 100.0%; Pred. No. 6.4e-177; Indels 0; Gaps 0;  
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 60  
Db 9 MAAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 68  
QY 61 SDPPCVRGMDSCSLEIANWRHNETLKYQDFAFYGERSTOODIYAGSVQPIRLHLLGQNA 120  
Db 69 SDPPCVRGMDSCSLEIANWRHNETLKYQDFAFYGERSTOODIYAGSVQPIRLHLLGQNA 128  
QY 121 ASVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 180  
Db 129 ASVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIY 188  
QY 181 QEKVLDLLDPASGLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLN 240  
Db 189 QEKVLDLLDPASGLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLN 248  
QY 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGEEDNRRTGNKGLRKESGAINTS 300  
Db 249 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGEEDNRRTGNKGLRKESGAINTS 308  
QY 301 LFVLGKVVVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 360  
Db 309 LFVLGKVVVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALN 368  
QY 361 FAARSKEVIN 370  
Db 369 FAARSKEVIN 378

RESULT 3  
US-10-797-893-4  
; Sequence 4, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/797,893  
; PRIOR FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Human  
US-10-797-893-4  
Query Match 99.7%; Score 1873; DB 16; Length 512;  
Best Local Similarity 100.0%; Pred. No. 2.2e-176; Indels 0; Gaps 0;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 61  
Db 3 AAGGSTQORREMAAASAAAIISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 62  
QY 62 DPCCVRGMDSCSLEIANWRHNETLKYQDFAFYGERSTOODIYAGSVQPIRLHLLGQNA 121  
Db 63 DPCCVRGMDSCSLEIANWRHNETLKYQDFAFYGERSTOODIYAGSVQPIRLHLLGQNA 122  
QY 122 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 181  
Db 123 SVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182  
QY 182 EKVLDLLDPASGLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLN 241  
Db 183 EKVLDLLDPASGLVIREDCRGNILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLN 242  
QY 242 RSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGEEDNRRTGNKGLRKESGAINTS 301  
Db 243 RSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGEEDNRRTGNKGLRKESGAINTS 302  
QY 302 FVLGKVVVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALNF 361  
Db 303 FVLGKVVVDALNOGLPRVPYRDSKLTLLQDSLGSSAHSILIANIAPERFYLDTVSALNF 362  
QY 362 AARSKEVIN 370  
Db 363 AARSKEVIN 371  
RESULT 4  
US-10-797-893-6  
; Sequence 6, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/797,893  
; PRIOR FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20

```

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-6

Query Match
Best Local Similarity 93.9%; Score 1763; DB 16; Length 346;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 85
DB 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61

QY 86 LKQYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 145
DB 62 LKQYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 121

QY 146 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205
DB 122 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181

QY 206 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 265
DB 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 241

QY 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 325
DB 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301

QY 326 TRLLODLSGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 370
DB 302 TRLLODLSGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346

RESULT 5
US-09-925-300-1228
; Sequence 1228, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1228
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (147)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (435)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1228

Query Match
Best Local Similarity 93.9%; Score 1763; DB 9; Length 460;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 24 GAGCRLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 83
DB 2 GAGCRLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61

QY 84 ETLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 143
DB 62 ETLKYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 121

QY 144 PGVPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRQ 203
DB 122 PGVPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRQ 181

QY 204 NILPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 263
DB 182 NILPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPR 241

QY 264 RQREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDS 323
DB 242 RQREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDS 301

QY 324 KLTRLLODLSGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 370
DB 302 KLTRLLODLSGSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 348

RESULT 6
US-10-797-893-8
; Sequence 8, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-8

Query Match
Best Local Similarity 93.9%; Score 1763; DB 16; Length 487;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 85
DB 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61

QY 86 LKQYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 145
DB 62 LKQYQDAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQ 121

QY 146 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205
DB 122 VIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181

QY 206 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 265
DB 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ 241

QY 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 325
DB 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301

```

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159031
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(377)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58448C.1.pep
; US-10-437-963-159031

Query Match      31.8%; Score 597.5; DB 16; Length 377;
Best Local Similarity 39.7%; Pred. No. 3.7e-50;
Matches 138; Conservative 69; Mismatches 112; Indels 29; Gaps 8;

QY 30 LSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANNRNHOETLK- 87
Db 30 LSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANNRNHOETLK- 87
QY 1 MATAAATOSQP---VRVVLVRVPHLPSEANGAACPVLGLGSHPGGEVTVQLKDQYTSRN 57
Db 1 MATAAATOSQP---VRVVLVRVPHLPSEANGAACPVLGLGSHPGGEVTVQLKDQYTSRN 57
QY 88 --YQFDYFGERSTQDDIYAGSVQPIRLHLLGONASVLAAYGPTGAGKTHMLGSPQPG 145
Db 88 --YQFDYFGERSTQDDIYAGSVQPIRLHLLGONASVLAAYGPTGAGKTHMLGSPQPG 145
QY 58 ECKYKLDAPFGQESRVEIFDQESAVIFGIFEGTNATVAYGATGSGKTYTMQGTEDLPG 117
Db 58 ECKYKLDAPFGQESRVEIFDQESAVIFGIFEGTNATVAYGATGSGKTYTMQGTEDLPG 117
QY 146 VIPRALMDLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 205
Db 146 VIPRALMDLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 205
QY 118 LMKSTVLAICTGT-----WC-SVEISYEVYVMERCYDLLEPKAREIMVLDKGNL 167
Db 118 LMKSTVLAICTGT-----WC-SVEISYEVYVMERCYDLLEPKAREIMVLDKGNL 167
QY 206 LIPGLSQKPISSFADFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
Db 206 LIPGLSQKPISSFADFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
QY 168 QLKGLAWVPVRSLEBFHEIYSGVQRKVAHTGLNDVSSRSHAVLSIRTT-----DV 220
Db 168 QLKGLAWVPVRSLEBFHEIYSGVQRKVAHTGLNDVSSRSHAVLSIRTT-----DV 220
QY 266 REGKLYI---DLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQLPRVPYRD 322
Db 266 REGKLYI---DLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQLPRVPYRD 322
QY 221 VEGKMLNITSXDLAGNEDNRRTCNIGIRLOESAKINQSLFALSNNVISALNKKEPRIPYRE 280
Db 221 VEGKMLNITSXDLAGNEDNRRTCNIGIRLOESAKINQSLFALSNNVISALNKKEPRIPYRE 280
QY 323 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370
Db 323 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370
QY 281 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370
Db 281 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370

RESULT 9
US-10-108-260A-2692
; Sequence 2692, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cdna
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2692
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-108-260A-2692

Query Match      29.8%; Score 559; DB 15; Length 548;
Best Local Similarity 38.9%; Pred. No. 4.2e-46;
Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY 43 RRVAVRLRPFP---VDGTAGAS-----DPCVGRGMDSCSLEIANRNHOETLK-YQFADF 93
Db 43 RRVAVRLRPFP---VDGTAGAS-----DPCVGRGMDSCSLEIANRNHOETLK-YQFADF 93
QY 11 QLMVALVRPISVAELBEGATLIAHKVDEQVMVMDPMDPDILLRAHRSREKSYLFDVA 70
Db 11 QLMVALVRPISVAELBEGATLIAHKVDEQVMVMDPMDPDILLRAHRSREKSYLFDVA 70

```

```

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159031
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(377)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58448C.1.pep
; US-10-437-963-159031

Query Match      31.8%; Score 597.5; DB 16; Length 377;
Best Local Similarity 39.7%; Pred. No. 3.7e-50;
Matches 138; Conservative 69; Mismatches 112; Indels 29; Gaps 8;

QY 30 LSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANNRNHOETLK- 87
Db 30 LSKIGATRRPPPARVRVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANNRNHOETLK- 87
QY 1 MATAAATOSQP---VRVVLVRVPHLPSEANGAACPVLGLGSHPGGEVTVQLKDQYTSRN 57
Db 1 MATAAATOSQP---VRVVLVRVPHLPSEANGAACPVLGLGSHPGGEVTVQLKDQYTSRN 57
QY 88 --YQFDYFGERSTQDDIYAGSVQPIRLHLLGONASVLAAYGPTGAGKTHMLGSPQPG 145
Db 88 --YQFDYFGERSTQDDIYAGSVQPIRLHLLGONASVLAAYGPTGAGKTHMLGSPQPG 145
QY 58 ECKYKLDAPFGQESRVEIFDQESAVIFGIFEGTNATVAYGATGSGKTYTMQGTEDLPG 117
Db 58 ECKYKLDAPFGQESRVEIFDQESAVIFGIFEGTNATVAYGATGSGKTYTMQGTEDLPG 117
QY 146 VIPRALMDLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 205
Db 146 VIPRALMDLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCRNI 205
QY 118 LMKSTVLAICTGT-----WC-SVEISYEVYVMERCYDLLEPKAREIMVLDKGNL 167
Db 118 LMKSTVLAICTGT-----WC-SVEISYEVYVMERCYDLLEPKAREIMVLDKGNL 167
QY 206 LIPGLSQKPISSFADFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
Db 206 LIPGLSQKPISSFADFERHFLPASNRVTGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 265
QY 168 QLKGLAWVPVRSLEBFHEIYSGVQRKVAHTGLNDVSSRSHAVLSIRTT-----DV 220
Db 168 QLKGLAWVPVRSLEBFHEIYSGVQRKVAHTGLNDVSSRSHAVLSIRTT-----DV 220
QY 266 REGKLYI---DLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQLPRVPYRD 322
Db 266 REGKLYI---DLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQLPRVPYRD 322
QY 221 VEGKMLNITSXDLAGNEDNRRTCNIGIRLOESAKINQSLFALSNNVISALNKKEPRIPYRE 280
Db 221 VEGKMLNITSXDLAGNEDNRRTCNIGIRLOESAKINQSLFALSNNVISALNKKEPRIPYRE 280
QY 323 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370
Db 323 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370
QY 281 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370
Db 281 SKLTRLODSLGSAHSILIANIAPERFYLDVTSALNFAARKEVIN 370

RESULT 9
US-10-108-260A-2692
; Sequence 2692, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cdna
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2692
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-108-260A-2692

Query Match      29.8%; Score 559; DB 15; Length 548;
Best Local Similarity 38.9%; Pred. No. 4.2e-46;
Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY 43 RRVAVRLRPFP---VDGTAGAS-----DPCVGRGMDSCSLEIANRNHOETLK-YQFADF 93
Db 43 RRVAVRLRPFP---VDGTAGAS-----DPCVGRGMDSCSLEIANRNHOETLK-YQFADF 93
QY 11 QLMVALVRPISVAELBEGATLIAHKVDEQVMVMDPMDPDILLRAHRSREKSYLFDVA 70
Db 11 QLMVALVRPISVAELBEGATLIAHKVDEQVMVMDPMDPDILLRAHRSREKSYLFDVA 70

```



```

366 GTVFCYGTAGAGKTYTLMGTGMENPGVWVLAINDLFSKVTKQH-----SIKLSYLEIY 418
181 QEKVLDDLPASGDVLVIREDCRGNILIPGLSOKPISSFADPFERHFLPASRNRTVGCATRLN 240
419 NETVRDLLSPGS-PLNLRDQKG-IVAAGUTQSVVYSTDEVMELLQKGNKTRTTEPTFRVN 476
241 QRSRSRSHAVLLVKVDQORERLAPPRQREGKUYLIDLAGSEDNRRRTGNKGLRKESGAINTS 300
477 ETSRSRSHAVLQVVVEYRSLDGVNVKQAGKSLIDLAGSERALATDQRTQRSIEGANINRS 536
301 LFLVGLKVDNALNOGLPRVPYRDSKLTRLIDSLGGGSAHSILIANIAPERRFYLDVTSAALN 360
537 LLALSSCINALVEGKHHPYRNSKLTQLLKDSLGGACNTVMIANISFNLSPGQTQTLH 596
361 FAARSKEV 368
597 WADRAKEI 604

RESULT 12
US-10-437-963-103943
; Sequence 103943, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

```

APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 103943  
LENGTH: 776  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_101324C.1.pap  
US-10-437-963-103943

Query Match 29.0%; Score 544.5; DB 16; Length 776;  
Best Local Similarity 35.3%; Pred. No. 1.9e-44; Indels 29; Gaps 6;  
Matches 133; Conservative 71; Mismatches 144;

QY 1 MAAGSTQORREMAAASAAAGAGRCRLSKIGATRRPPARVVRVAVRLRPFVDGTAGA 60  
DB 178 LSLGGMAARLKTAGEAGAGNGDAAG-----SRIMVFLRPMRSKEKDA 222  
QY 61 SDPPCVRGMDSCSLEIANWRNHQETLK-----YQDFAYGERSTQODIYAGSVQPI 112  
DB 223 GSRSCVKIWNKDVVTEFASETDYLRKLRVGRSHFCFDSFPDTTQAEVISTTSDLV 282  
QY 113 RHLEGNASVLYAGTCACTHMLGSPQGVIPRALMDLLQLTREEGAGCRPWALS 172  
DB 283 EGVLRNGTVFCYATGAGTGYTLMGTWESPGVWVLAIKDLFTKVRQSHDGNH---SI 339  
QY 173 TMSYLEIYQEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNR 232  
DB 340 QLSYLEVNETVRLDLSRGR-PLLRREDKQGTV-AAGLTHYRAYSTDEVKMLQOQGNR 397  
QY 233 TVGATRLNQRSSRHAVLLVKVDQRE-RLAPPROREGKLYLIDLAGSDNRRTGNKGLRL 291  
DB 398 TTEFRVNETSRSHAILQVIEYRSIDGSGIVTRVGLSLIDLAGSERALATQRTQS 457  
QY 292 KESGAINTSFLVGLKVDALNQLPRVPYRDSKLTRELQDSLGGSAAHSILIANIAPERRF 351  
DB 458 IEGANINRSLALSSCINALVEGKKHPIYRNSKLTQLLKDSLGSGCNTWMIANISPSNLS 517  
QY 352 YLDTVSALNFAARSKEV 368  
DB 518 FGETQNTLHWADRAKEI 534

RESULT 13  
US-09-883-096-5  
Sequence 5, Application US/09883096  
Patent No. US20020110883A1  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Craven, Andrew  
APPLICANT: Yu, Ming  
APPLICANT: Sakowicz, Roman  
APPLICANT: Patel, Umesh A.  
APPLICANT: Davies, Katherine A.  
TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE  
FILE REFERENCE: 020552-001410US  
CURRENT APPLICATION NUMBER: US/09/883,096  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 09/594,655  
PRIOR FILING DATE: 2000-06-15  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1

APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 103943  
LENGTH: 776  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_101324C.1.pap  
US-10-437-963-103943

Query Match 29.0%; Score 544.5; DB 16; Length 776;  
Best Local Similarity 35.3%; Pred. No. 1.9e-44; Indels 29; Gaps 6;  
Matches 133; Conservative 71; Mismatches 144;

QY 1 MAAGSTQORREMAAASAAAGAGRCRLSKIGATRRPPARVVRVAVRLRPFVDGTAGA 60  
DB 178 LSLGGMAARLKTAGEAGAGNGDAAG-----SRIMVFLRPMRSKEKDA 222  
QY 61 SDPPCVRGMDSCSLEIANWRNHQETLK-----YQDFAYGERSTQODIYAGSVQPI 112  
DB 223 GSRSCVKIWNKDVVTEFASETDYLRKLRVGRSHFCFDSFPDTTQAEVISTTSDLV 282  
QY 113 RHLEGNASVLYAGTCACTHMLGSPQGVIPRALMDLLQLTREEGAGCRPWALS 172  
DB 283 EGVLRNGTVFCYATGAGTGYTLMGTWESPGVWVLAIKDLFTKVRQSHDGNH---SI 339  
QY 173 TMSYLEIYQEKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNR 232  
DB 340 QLSYLEVNETVRLDLSRGR-PLLRREDKQGTV-AAGLTHYRAYSTDEVKMLQOQGNR 397  
QY 233 TVGATRLNQRSSRHAVLLVKVDQRE-RLAPPROREGKLYLIDLAGSDNRRTGNKGLRL 291  
DB 398 TTEFRVNETSRSHAILQVIEYRSIDGSGIVTRVGLSLIDLAGSERALATQRTQS 457  
QY 292 KESGAINTSFLVGLKVDALNQLPRVPYRDSKLTRELQDSLGGSAAHSILIANIAPERRF 351  
DB 458 IEGANINRSLALSSCINALVEGKKHPIYRNSKLTQLLKDSLGSGCNTWMIANISPSNLS 517  
QY 352 YLDTVSALNFAARSKEV 368  
DB 518 FGETQNTLHWADRAKEI 534

RESULT 13  
US-09-883-096-5  
Sequence 5, Application US/09883096  
Patent No. US20020110883A1  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Craven, Andrew  
APPLICANT: Yu, Ming  
APPLICANT: Sakowicz, Roman  
APPLICANT: Patel, Umesh A.  
APPLICANT: Davies, Katherine A.  
TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE  
FILE REFERENCE: 020552-001410US  
CURRENT APPLICATION NUMBER: US/09/883,096  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 09/594,655  
PRIOR FILING DATE: 2000-06-15  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5  
LENGTH: 383  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: HsKip3a  
OTHER INFORMATION: fragment  
OTHER INFORMATION: Amino acid sequence of HsKip3a fragment used in  
OTHER INFORMATION: the ATPase assay (Figure 4).  
US-09-883-096-5

Query Match 28.1%; Score 528; DB 9; Length 383;  
Best Local Similarity 38.7%; Pred. No. 2.9e-43;  
Matches 138; Conservative 46; Mismatches 127; Indels 46; Gaps 8;

QY 38 RPPPARVRVAVRLRPFV-----DG-----TAGASDPVCVRGMDSCSLEI 76  
DB 15 RPPTPRELDSQR-RPVQVQVDERVLVFNPEEPDFGLKKGWGTGDKKKGKD----- 67  
QY 77 ANWRNHQETLKQVDFADFYGERSTQODIYAGSVQPIRLHLEGNASVLYAGTCACTH 136  
DB 68 -----LTFVDRVFEAATQODVFOHTHSDVLSFLOGYNCSVFAYGATGAKTHT 118  
QY 137 MLGSPQGVIPRALMDLLQLTREEGAGCRPWALSVTMSYLEIYQEKVLDLDPASGDLV 196  
DB 119 MLGREGDPGIM---YLTTVELYRLEARQEKHFELISYQEVYNEQIHDLEP-KGFLA 174  
QY 197 IREDCRGNILIPGLSQKPISSPADFERHFLPASRNRVTGATRLNQRSSRHAVLLVKVDQ 256  
DB 175 IREDPDKGVVQGLSFHPQASAEQLLEILTRGNRNRTOHPTDANATSSRSHAIPOIFVKQ 234  
QY 257 REELAPPRO--REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVGLKVDALNQG 314  
DB 235 QRVPGLTQAVQVAXMSLIDLAGSERASSTHAKGERLREGANINRSLALINVLNALADA 294  
QY 315 LPR---VPYRDSKLTRELQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEV 368  
DB 295 KGRKTHVPYRDSKLTRELQDSLGSGCNTWMIANISPSLYTETDNTYTLADRAKEI 351

RESULT 14  
US-09-883-096-2  
Sequence 2, Application US/09883096  
Patent No. US20020110883A1  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Craven, Andrew  
APPLICANT: Yu, Ming  
APPLICANT: Sakowicz, Roman  
APPLICANT: Patel, Umesh A.  
APPLICANT: Davies, Katherine A.  
TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE  
FILE REFERENCE: 020552-001410US  
CURRENT APPLICATION NUMBER: US/09/883,096  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 09/594,655  
PRIOR FILING DATE: 2000-06-15  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 864  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino acid sequence encoded by human kinesin motor  
OTHER INFORMATION: protein gene HsKip3a (Figure 1).  
OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
OTHER INFORMATION: sequence of HsKip3a.  
US-09-883-096-2

Query Match 28.1%; Score 528; DB 9; Length 864;  
Best Local Similarity 38.7%; Pred. No. 9.7e-43;  
Matches 138; Conservative 46; Mismatches 127; Indels 46; Gaps 8;



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 58.709 Seconds  
(without alignments)  
2081.726 Million cell updates/sec

Title: US-10-797-893-6  
Perfect score: 1768  
Sequence: 1 MGRCLSKIGATRRPPARV.....FYLDTSALNFAARSKEVIN 346

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1768	100.0	346	16 US-10-797-893-6	Sequence 6, Appli
2	1768	100.0	487	16 US-10-797-893-8	Sequence 8, Appli
3	1763	99.7	370	16 US-10-797-893-2	Sequence 2, Appli
4	1763	99.7	490	15 US-10-334-143-80	Sequence 80, Appli
5	1763	99.7	512	16 US-10-797-893-4	Sequence 4, Appli
6	1753	99.2	460	9 US-09-925-300-1228	Sequence 1228, Ap
7	597.5	33.8	377	16 US-10-437-963-159031	Sequence 159031,
8	593.5	33.6	589	17 US-10-425-115-320209	Sequence 320209,
9	559	31.6	548	15 US-10-108-260A-2692	Sequence 2692, Ap
10	550	31.1	783	17 US-10-425-115-332919	Sequence 332919,
11	550	31.1	854	15 US-10-425-114-59708	Sequence 59708, A
12	537	30.4	776	16 US-10-437-963-103943	Sequence 103943,
13	528	29.9	383	9 US-09-883-096-5	Sequence 5, Appli

14	528	29.9	864	9 US-09-883-096-2	Sequence 2, Appli
15	527	29.8	757	15 US-10-220-120-366	Sequence 366, App
16	527	29.8	757	16 US-10-363-829-316	Sequence 316, App
17	526	29.8	375	15 US-10-332-089-4	Sequence 4, Appli
18	526	29.8	409	15 US-10-332-089-6	Sequence 6, Appli
19	526	29.8	774	14 US-10-369-493-2361	Sequence 2361, Ap
20	526	29.8	1388	14 US-10-146-473-82	Sequence 82, Appli
21	526	29.8	1388	15 US-10-173-999-32	Sequence 32, Appli
22	526	29.8	1388	15 US-10-332-089-2	Sequence 2, Appli
23	526	29.8	1388	15 US-10-188-832-164	Sequence 164, App
24	523.5	29.6	834	15 US-10-287-226-306	Sequence 306, App
25	521.5	29.5	1401	15 US-10-287-226-142	Sequence 142, App
26	519.5	29.4	1826	16 US-10-408-765A-1477	Sequence 1477, Ap
27	518.5	29.3	1382	16 US-10-437-963-176714	Sequence 176714,
28	518	29.3	1115	15 US-10-287-226-302	Sequence 302, App
29	518	29.3	1232	14 US-10-116-712-670	Sequence 670, App
30	518	29.3	1232	16 US-10-408-765A-2153	Sequence 2153, Ap
31	518	29.3	1235	15 US-10-334-143-8	Sequence 8, Appli
32	517.5	29.3	1001	17 US-10-425-115-231231	Sequence 231231,
33	517.5	29.3	1324	15 US-10-287-226-314	Sequence 314, App
34	516	29.2	952	17 US-10-739-930-7903	Sequence 7903, Ap
35	515	29.1	1030	15 US-10-425-114-62748	Sequence 62748, A
36	513.5	29.0	805	14 US-10-369-493-21903	Sequence 21903, A
37	513	29.0	1237	15 US-10-334-143-33	Sequence 33, Appli
38	510	28.8	672	16 US-10-408-765A-1664	Sequence 1664, Ap
39	509	28.8	1103	9 US-09-847-874A-1	Sequence 1, Appli
40	509	28.8	1103	14 US-10-458-162-1	Sequence 1, Appli
41	507	28.7	928	14 US-10-080-608A-23	Sequence 23, Appli
42	507	28.7	928	14 US-10-370-685-112	Sequence 112, App
43	507	28.7	1232	14 US-10-116-712-664	Sequence 664, App
44	507	28.7	1232	14 US-10-116-712-669	Sequence 669, App
45	504	28.5	338	9 US-09-883-096-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-797-893-6  
; Sequence 6, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Human  
; US-10-797-893-6

Query Match	100.0%;	Score 1768;	DB 16;	Length 346;
Best Local Similarity	100.0%;	Pred. No. 1.7e-168;		
Matches 346;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQE	60	
Db	1	MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQE	60	
QY	61	TLKQYDFAFYGERSTQDDIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHTMLGSPQ	120	
Db	61	TLKQYDFAFYGERSTQDDIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHTMLGSPQ	120	
QY	121	GVIPRALMDLLQLTREGEAGRPWALSVTMSYLEIYQKVLDDLPASGDLVITREDCRGN	180	

Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
 QY 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240  
 Db 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240  
 QY 241 QREGKLYLIDLAGESEDRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300  
 Db 241 QREGKLYLIDLAGESEDRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 Db 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 2

US-10-797-893-8  
 ; Sequence 8, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; PRIOR FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 487  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-10-797-893-8

Query Match 100.0%; Score 1768; DB 16; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-168;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 60  
 Db 1 MGRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 60  
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAGPTGAGKTHMLGSPEOP 120  
 Db 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAGPTGAGKTHMLGSPEOP 120  
 QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
 Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
 QY 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240  
 Db 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240  
 QY 241 QREGKLYLIDLAGESEDRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300  
 Db 241 QREGKLYLIDLAGESEDRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 Db 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 3

US-10-797-893-2  
 ; Sequence 2, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; PRIOR FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 370  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-10-797-893-2

Query Match 99.7%; Score 1763; DB 16; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-168;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 61  
 Db 26 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOE 85  
 QY 62 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAGPTGAGKTHMLGSPEOP 121  
 Db 86 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAAGPTGAGKTHMLGSPEOP 145  
 QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181  
 Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205  
 QY 182 LIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 241  
 Db 206 LIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSSRSHAVLLVKVDQERLAPFR 265  
 QY 242 REGKLYLIDLAGESEDRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSKL 301  
 Db 266 REGKLYLIDLAGESEDRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSKL 325  
 QY 302 TRLLODSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 Db 326 TRLLODSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

RESULT 4

US-10-334-143-80  
 ; Sequence 80, Application US/10334143  
 ; Publication No. US20040009549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
 ; APPLICANT: SUDARSANAM, SUCHA  
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
 ; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
 ; FILE REFERENCE: 038602/1543  
 ; CURRENT APPLICATION NUMBER: US/10/334,143  
 ; CURRENT FILING DATE: 2002-12-31  
 ; PRIOR APPLICATION NUMBER: 60/343,169  
 ; PRIOR FILING DATE: 2001-12-31  
 ; NUMBER OF SEQ ID NOS: 207  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 80  
 ; LENGTH: 490  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-334-143-80

Query Match 99.7%; Score 1763; DB 15; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 9e-168;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDTGASDPPPCVRGMDSCSLEIANWRNHQET 61  
 DB 34 GRCLSKIGATRRPPPARVRVAVRLRPFVDTGASDPPPCVRGMDSCSLEIANWRNHQET 93  
 QY 62 LKQFDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121  
 DB 94 LKQFDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 153  
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181  
 DB 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 213  
 QY 182 LIPGLSKQPISSFADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241  
 DB 214 LIPGLSKQPISSFADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 273  
 QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301  
 DB 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 333  
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 346  
 DB 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 378

RESULT 5  
 US-10-797-893-4  
 ; Sequence 4, Application US/10797893  
 ; Publication No. US2004014237A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 512  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-10-797-893-4

Query Match 99.7%; Score 1763; DB 16; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 9.6e-168;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDTGASDPPPCVRGMDSCSLEIANWRNHQET 61  
 DB 27 GRCLSKIGATRRPPPARVRVAVRLRPFVDTGASDPPPCVRGMDSCSLEIANWRNHQET 86  
 QY 62 LKQFDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121  
 DB 87 LKQFDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 146  
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181  
 DB 147 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 206  
 QY 182 LIPGLSKQPISSFADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241  
 DB 207 LIPGLSKQPISSFADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 266  
 QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301  
 DB 267 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 326  
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 346

DB 327 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 371  
 RESULT 6  
 US-09-925-300-1228  
 ; Sequence 1228, Application US/09925300  
 ; Patent No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1228  
 ; LENGTH: 460  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (75)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (147)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (435)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-925-300-1228

Query Match 99.2%; Score 1753; DB 9; Length 460;  
 Best Local Similarity 99.4%; Pred. No. 8.2e-167;  
 Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDTGASDPPPCVRGMDSCSLEIANWRNHQET 61  
 DB 4 GRCLSKIGATRRPPPARVRVAVRLRPFVDTGASDPPPCVRGMDSCSLEIANWRNHQET 63  
 QY 62 LKQFDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121  
 DB 64 LKQFDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 123  
 QY 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181  
 DB 124 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 183  
 QY 182 LIPGLSKQPISSFADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241  
 DB 184 LIPGLSKQPISSFADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQERLAPFRQ 243  
 QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 301  
 DB 244 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL 303  
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 346  
 DB 304 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVIN 348

RESULT 7  
 US-10-437-963-159031  
 ; Sequence 159031, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 159031  
LENGTH: 377  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(377)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_58448C.1.psp  
US-10-437-963-159031

Query Match 33.8%; Score 597.5; DB 16; Length 377;  
Best Local Similarity 39.7%; Pred. No. 6.5e-51;  
Matches 138; Conservative 69; Mismatches 112; Indels 29; Gaps 8;  
QY 6 LSKIGATRRPPARVAVRLRPFDGTAGADPPCVRGMDSCSLEIANRNHQTLLK- 63  
DB 1 MATAAATQSQP---VRVLRVRLPSEANSSEAACVGLGSHPGGVTVQLKDYTSRN 57  
QY 64 --YQDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121  
DB 58 ECVKLDAPFGQSSRVCEIFDQVSAVPGIFGTNATFAYGATGSGKTYTMQGTEDPLPG 117  
QY 122 VTPRALMDLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGLDIVREDCRNI 181  
DB 118 LMXSTVIALCTGT-----WC-SVEISYEVYMERCYDLEPRKEIMVLDKDGNL 167  
QY 182 LTPGLSQPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQRLAPFRQ 241  
DB 168 QLKGLAWVPVRSLEEFHEIYSIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 220  
QY 242 REGKLYLI---DLAGSEDNRTGNKGLKESGAINTSFLVLGKVVDALNQGLPRVPYRD 298  
DB 221 VKGKMLITSXDLAGNEDNRTCNIGIRLQESAKINQSLFALSNNVISALNKKEPRIPYRE 280  
QY 299 SKLTRILOSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
DB 281 SKLTRILOSLGNSHAVMIACLNPVE--YQAVHTVSLAARSHVTN 326

RESULT 8  
US-10-425-115-320209  
Sequence 320209, Application US/10425115  
Publication No. US2004021427A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 320209  
LENGTH: 589  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:

OTHER INFORMATION: Clone ID: MRT4577\_55097C.1.psp  
US-10-425-115-320209  
Query Match 33.6%; Score 593.5; DB 17; Length 589;  
Best Local Similarity 39.2%; Pred. No. 3.2e-50;  
Matches 135; Conservative 71; Mismatches 107; Indels 31; Gaps 8;  
QY 17 PAR-----VRVAVRLRPFDGTAGASDPPCVRGMDSC-----SLEIANRNHQTLL 62  
DB 7 PARSSAHLSPQVRVLRVRLPFLSEASATAPCV-SLLGCHPGGVTVQKLD-QHTSRSE 64  
QY 63 KYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 122  
DB 65 QYKLDAPFRQEDSVSQIPDQEVRAVIPSIFEGINATVFAYGATGSGKTYTMQGTEDPFL 124  
QY 123 IPRALMDLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGLDIVREDCRNI 182  
DB 125 IPLAASTILALCT-----GTWC-SVEISYEVYMERCYDLEPRKEIMVLDKDGNM 177  
QY 183 IPGLSQPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQRLAPFRQ 242  
DB 178 LKGLSWVPVRSMEEFQELYSIGVQRKVAHTGLNDVSSRSHAVLSRVSSTDVV----- 231  
QY 243 EGKLYLIDLAGSEDNRTGNKGLKESGAINTSFLVLGKVVDALNQGLPRVPYRD SKLT 302  
DB 232 KGKLNLDLAGSEDNKLTNEGIRLQESSKINQSLFALSNNVISALNKNEHRIPYRQSKLT 291  
QY 303 RLLODSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
DB 292 RLLRSLGSGSRTVMIACLNPAAE--YQESANTVSLAARSCHIEN 333  
RESULT 9  
US-10-108-260A-2692  
Sequence 2692, Application US/10108260A  
Publication No. US20040005560A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
FILE REFERENCE: HI-A0106  
CURRENT APPLICATION NUMBER: US/10/108, 260A  
CURRENT FILING DATE: 2002-03-27  
NUMBER OF SEQ ID NOS: 5458  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2692  
LENGTH: 548  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-108-260A-2692  
Query Match 31.6%; Score 559; DB 15; Length 548;  
Best Local Similarity 38.9%; Pred. No. 8.3e-47;  
Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;  
QY 19 RRVAVRLRPF--VDGTAGAS-----DPPCVRGMDSCSLEIANRNHQTLLK-YQDFAF 69  
DB 11 QLMVALRVPI-SVAELEBEGATLIAHKVDEQMVLMDEPDIDILRAHRSREKSYLFDVA 70  
QY 70 YGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPGVPRALMD 129  
DB 71 FDFATQEMVYQATTKSLIEGVISYNATVFAYGTGCGKTYTMGLTQDEPGIYVQTLLND 130  
QY 130 LLOLTREBAGRPWALSVTMSYLEIYQEKVLDLDPASGLDIVREDCRNIPLGSLQK 189  
DB 131 LFRATEETSD---MEYEVMSYLEIYNEMRLDNLPSLGYLELREDSKGVIVAGITEV 187  
QY 190 PISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQRLAPFRQ--REGKLY 247  
DB 188 STINAKEIMQLMKGNRQRTQPTAANTSSRSHAVLQVTVRQSRVKNILQEVQRGLF 247  
QY 248 LIDLAGSEDNRTGNKGLKESGAINTSFLVLGKVVDAL--NOGLPR--VPYRDSKLTLL 305  
DB 248 MIDLAGSERASQTNRGQRMKEGAHNRSLALGNCINALSDKSGNKYINRYRDSKLTLL 307



Qy	306 QBSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEV 344   :       : :   :   : : :   :   :   :
Db	308 KDSLGGNSRTVMIAHTSPASSAFEESRNTLTYYAGRAKNI 346   :       : :   :   : : :   :   :   :

## RESULT, T 10

```

US-10-425-115-332919
Sequence 332919, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21 (53222) B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 332919
LENGTH: 783
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1) : (783)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MFT4577_66737C.1.pgp

```

Query Match	31.1%;	Score 550;	DB 17;	Length 783;
Best Local Similarity	38.2%;	Pred. No. 1.1e-45;		
Matches 132;	Conservative	68;	Mismatches 128;	Indels 18;
				Gaps 5;

[illegible]

## RESULT 11

US-10-425-114-59708  
; Sequence 59708, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei

```

: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
:
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
:
: FILE REFERENCE: 38-21(53313)B
:
: CURRENT APPLICATION NUMBER: US/10/425,114
:
: CURRENT FILING DATE: 2003-04-28
:
: NUMBER OF SEQ ID NOS: 73128
:
: SEQ ID NO 59708
:
: LENGTH: 854
:
: TYPE: PRT
:
: ORGANISM: Zea mays
:
: FEATURE:
:
: OTHER INFORMATION: Clone ID: LIB3689-256-H1_FLI pep
US-10-425-114-59708

```

Query Match 31.1%; Score 550; DB 15; Length 854;

Best Local Similarity 38.2%; Score 3507; Indels 45; Gaps 634;  
Matches 132; Conservative 68; Mismatches 128; Indels 18; Gaps 5;

```
QY      8 KIGA-TRRPPPAVRVAVLRPPFDGTAGASDPPCVGRMDSCSLEIANRNHQTLLK--- 63
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     268 EVGAGTAATAAESRLVFRVRPMRSRKEKEAGSSCVKIVNRKEVFLTESASENDYLRKR 327
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
QY     64 -----YQDAFYGERSTQQDIYAGSVQPIRLHLELLEGONASVLA YGPTGAGKTHMLGSPE 118
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     328 GRDSHCFDSVPFDSITQAEVYSTADLVEGVLOQENGTVFCYGATGAKTYTLMGTME 387
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
QY     119 QGVIVIPRALMDLLQLTREAGEGRPWALSVTMSYLEIYQEKVLDDPASPAGDLVIREDCR 178
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     388 NPGVMVLAINDLFSKYTKNH-----SIKLSYLEIYNTVDDLSPGS-PLNLRDQK 439
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
QY     179 GNILIPGLGQKPTSSPADFERHFLPARNRTVGATRLNQRSSHVA LLVKVQDQERLAP 238
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     440 G-IVAAGLTQRSVYSTDEVMELLQKGNKRTTETPRVNETSSRSHAVLQVVVEVRSLDGV 498
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
QY     239 FRQREKGLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGKVVDA LNQGLPRVPYRD 298
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     499 NVKRAGKLSLIDLAGSERALATDQRTQSTEGANINRSLLALSSCINALVEGKKHIPYRN 558
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
QY     299 SKULTRLDQSLGSSAHSILIANAPERRFYLDTVTSALNFAARKEV 344
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db     559 SKITQLLKDSLQAGACNTVMIANISPNLSFGFETQNTLHWADRAKEI 604
      :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
```

## RESULT 12

```

US-10-437-963-103943
; Sequence 103943, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)/B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 103943
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101324C.1.pep
US-10-437-963-103943

```

Query Match	30.4%	Score 537	DB 16	Length 776
Best Local Similarity	37.5%	Pred. No. 2.3e-44		
Matches 126; Conservative	68	Mismatches 128	Indels 1	





This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 13.5172 Seconds  
(without alignments)  
1697.543 Million cell updates/sec

Title: US-10-797-893-6  
Perfect score: 1768  
Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARKEVIN 346

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1768	100.0	346	3	US-09-724-224-6
2	1768	100.0	346	4	US-10-093-317-6
3	1768	100.0	487	3	US-09-724-224-8
4	1768	100.0	487	4	US-10-093-317-8
5	1763	99.7	370	3	US-09-724-224-2
6	1763	99.7	370	4	US-10-093-317-2
7	1763	99.7	512	3	US-09-724-224-4
8	1763	99.7	512	4	US-10-093-317-4
9	1763	99.7	665	3	US-09-595-684B-35
10	549	31.1	355	3	US-09-724-511-4
11	549	31.1	355	4	US-09-723-097-4
12	549	31.1	355	4	US-09-632-344-4
13	549	31.1	367	3	US-09-724-511-2
14	549	31.1	367	4	US-09-723-097-2
15	549	31.1	367	4	US-09-632-344-2
16	528	29.9	383	4	US-09-883-096-5
17	528	29.9	864	4	US-09-883-096-2
18	526	29.8	375	3	US-09-572-191-4
19	526	29.8	375	3	US-09-723-262-4
20	526	29.8	375	3	US-09-723-219-4
21	526	29.8	409	3	US-09-572-191-6
22	526	29.8	409	3	US-09-723-262-6
23	526	29.8	409	3	US-09-723-219-6
24	526	29.8	1388	3	US-09-572-191-2
25	526	29.8	1388	3	US-09-723-262-2
26	526	29.8	1388	3	US-09-723-219-2
27	521.5	29.5	522	4	US-09-592-054-4

28	521.5	29.5	1279	3	US-09-724-517-2	Sequence 2, Appli
29	521.5	29.5	1279	4	US-09-641-807A-2	Sequence 2, Appli
30	521.5	29.5	1279	4	US-09-723-096-2	Sequence 2, Appli
31	518.5	29.3	341	3	US-09-724-517-4	Sequence 4, Appli
32	518.5	29.3	341	4	US-09-641-807A-4	Sequence 4, Appli
33	518.5	29.3	341	4	US-09-723-096-4	Sequence 4, Appli
34	518	29.3	473	4	US-09-592-054-6	Sequence 6, Appli
35	517.5	29.3	342	3	US-09-641-806-2	Sequence 2, Appli
36	517.5	29.3	342	4	US-09-723-129-2	Sequence 2, Appli
37	517.5	29.3	342	4	US-09-722-862-2	Sequence 2, Appli
38	514.5	29.1	337	3	US-09-641-806-4	Sequence 4, Appli
39	514.5	29.1	337	4	US-09-723-129-4	Sequence 4, Appli
40	514.5	29.1	337	4	US-09-722-862-4	Sequence 4, Appli
41	514	29.1	1066	3	US-09-541-782-8	Sequence 8, Appli
42	514	29.1	1066	4	US-09-723-820-8	Sequence 8, Appli
43	514	29.1	1066	4	US-10-270-085-8	Sequence 8, Appli
44	514	29.1	1690	4	US-09-595-684B-39	Sequence 39, Appli
45	511.5	28.9	299	3	US-09-621-233-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-09-724-224-6  
; Sequence 6, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; FILE REFERENCE: their use  
; CURRENT APPLICATION NUMBER: US/09/724,224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-224-6

Query Match		100.0%	Score 1768	DB 3	Length 346
Best Local Similarity		100.0%	Pred. No. 1.2e-180		
Matches 346		Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	MGRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60		
Db	1	MGRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60		
Qy	61	TLKYQFADFYGERSTQDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHTMLGSPEQP	120		
Db	61	TLKYQFADFYGERSTQDIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHTMLGSPEQP	120		
Qy	121	GVTPRALMDLLOITREGAGRWALSVTMSYLEIQEKVLDLLDPASGDLVREDCRGN	180		
Db	121	GVTPRALMDLLOITREGAGRWALSVTMSYLEIQEKVLDLLDPASGDLVREDCRGN	180		
Qy	181	ILIPGLSQKPISSFADFPERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR	240		
Db	181	ILIPGLSQKPISSFADFPERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR	240		
Qy	241	QREGKLYLIDLAGSDNRRTGNKGLKESGAINTSFLVLGKVVDAALNQLPRVPYRDSK	300		
Db	241	QREGKLYLIDLAGSDNRRTGNKGLKESGAINTSFLVLGKVVDAALNQLPRVPYRDSK	300		
Qy	301	LTELLQDSLGGSASHSILIANIAPERFYLDTSALNFAARKEVIN	346		
Db	301	LTELLQDSLGGSASHSILIANIAPERFYLDTSALNFAARKEVIN	346		

```

RESULT 2
US-10-093-317-6
; Sequence 6, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-6

Query Match      100.0%; Score 1768; DB 4; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.2e-180;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB 1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAEGPTGAGKTHMTLGSPEQ 120
DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAEGPTGAGKTHMTLGSPEQ 120
QY 121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEIYQEKVLDLDPASGDLVIREDCRGN 180
DB 121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEIYQEKVLDLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQKPISSFADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPPR 240
DB 181 ILIPGLSQKPISSFADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPPR 240
QY 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVYRDSK 300
DB 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVYRDSK 300
QY 301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB 301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 4
US-10-093-317-8
; Sequence 8, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-10-093-317-8

Query Match      100.0%; Score 1768; DB 4; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.1e-180;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
DB 1 MGRCLSKIGATRRPPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60
QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAEGPTGAGKTHMTLGSPEQ 120
DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAEGPTGAGKTHMTLGSPEQ 120
QY 121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEIYQEKVLDLDPASGDLVIREDCRGN 180
DB 121 GVIPRALMDLLQLTREAGRGPWALSVMYSLEIYQEKVLDLDPASGDLVIREDCRGN 180
QY 181 ILIPGLSQKPISSFADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPPR 240
DB 181 ILIPGLSQKPISSFADFERHFLPASNRNRTVGATRLNQSRSSHAVLLVKVDQERLAPPR 240
QY 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVYRDSK 300
DB 241 QREGKLYLIDLAGESENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVYRDSK 300
QY 301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
DB 301 LTRLQDLSGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 3
US-09-724-224-8
; Sequence 8, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-8

Query Match      100.0%; Score 1768; DB 3; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.1e-180;

```

```
|||||
Db 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 5
US-09-724-224-2
; Sequence 2, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-2

Query Match 99.7%; Score 1763; DB 3; Length 370;
Best Local Similarity 100.0%; Pred. No. 4.5e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCCVGRGMDSCSLEIANWRNHQET 61
Db 26 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCCVGRGMDSCSLEIANWRNHQET 85

Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPG 121
Db 86 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPG 145

Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRGN 181
Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRGN 205

Qy 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241
Db 206 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVDQERLAPFRQ 265

Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 301
Db 266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 325

Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346
Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

RESULT 7
US-09-724-224-4
; Sequence 4, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
US-09-724-224-4

Query Match 99.7%; Score 1763; DB 3; Length 512;
Best Local Similarity 100.0%; Pred. No. 7.6e-180;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCCVGRGMDSCSLEIANWRNHQET 61
Db 27 GRCLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCCVGRGMDSCSLEIANWRNHQET 86

Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPG 121
Db 87 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLAGPTGAGKTHTMLGSPQPG 146

Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRGN 181
Db 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRGN 206

Qy 182 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVDQERLAPFRQ 241
Db 207 LIPGLSQKPISSFADFERHFLPASRNRVTGATRLNQSRSSHAVLLVKVDQERLAPFRQ 266

Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 301
```

Db 267 REGKYLIDLAGSDNRRTGNKGLRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 326  
Qy 302 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346  
Db 327 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 371

RESULT 8  
US-10-093-317-4  
; Sequence 4, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Human  
US-10-093-317-4

Query Match 99.7%; Score 1763; DB 4; Length 512;  
Best Local Similarity 100.0%; Pred. No. 7.6e-180;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVRLPPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
Db 27 GRCLSKIGATRRPPPARVRVAVRLPPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86  
Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPQPG 121  
Db 87 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPQPG 146  
Qy 122 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVIREDCRNI 181  
Db 147 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVIREDCRNI 206  
Qy 182 LIPGLSKQPISSFADFERRHFLPASRNTVGATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 241  
Db 207 LIPGLSKQPISSFADFERRHFLPASRNTVGATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 266  
Qy 242 REGKYLIDLAGSDNRRTGNKGLRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 301  
Db 267 REGKYLIDLAGSDNRRTGNKGLRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 326  
Qy 302 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346  
Db 327 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 371

RESULT 9  
US-09-595-684B-35  
; Sequence 35, Application US/09595684B  
; Patent No. 6544766  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Ohashi, Cara  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Vaisberg, Eugeni  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Yu, Ming  
; TITLE OF INVENTION: Human kinesins and methods of producing  
; and purifying human kinesins  
; FILE REFERENCE: cytop036  
; CURRENT APPLICATION NUMBER: US/09/595,684B

; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Human  
US-09-595-684B-35

Query Match 99.7%; Score 1763; DB 4; Length 665;  
Best Local Similarity 100.0%; Pred. No. 1.2e-179;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPARVRVAVRLPPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
Db 26 GRCLSKIGATRRPPPARVRVAVRLPPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85  
Qy 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPQPG 121  
Db 86 LKQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPQPG 145  
Qy 122 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVIREDCRNI 181  
Db 146 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDDLPASGDLVIREDCRNI 205  
Qy 182 LIPGLSKQPISSFADFERRHFLPASRNTVGATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 241  
Db 206 LIPGLSKQPISSFADFERRHFLPASRNTVGATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 265  
Qy 242 REGKYLIDLAGSDNRRTGNKGLRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 301  
Db 266 REGKYLIDLAGSDNRRTGNKGLRLKESGAINSTSLFVLGKVDALNOGLPRVPYRDSKL 325  
Qy 302 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 346  
Db 326 TRLLQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEVIN 370

RESULT 10  
US-09-724-511-4  
; Sequence 4, Application US/09724511  
; Patent No. 6391601  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6391601el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1046  
; CURRENT APPLICATION NUMBER: US/09/724,511  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 355  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-511-4

Query Match 31.1%; Score 549; DB 3; Length 355;  
Best Local Similarity 39.0%; Pred. No. 4.6e-50;  
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

Qy 20 VRVAVRLRP-----FVDGTAGASDPP-----PCVRGMDSCSLEIANWRNHQ 59  
Db 4 MKVYVVRREPTKEKAAAGFKVHVVDVKHILVDFPKQEVSTFHGKKTNNQVKKQN--- 61  
Qy 60 ETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPQPG 119  
Db 62 KDLKFFVDAVDFDETSTQSEVFHTTKPILRSFLNGYNTVLAAYGATGAGKTHMLGSADE 121

Query Match 31.1%; Score 549; DB 3; Length 355;  
Best Local Similarity 39.0%; Pred. No. 4.6e-50;  
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

Qy 20 VRVAVRLRP-----FVDGTAGASDPP-----PCVRGMDSCSLEIANWRNHQ 59  
Db 4 MKVYVVRREPTKEKAAAGFKVHVVDVKHILVDFPKQEVSTFHGKKTNNQVKKQN--- 61  
Qy 60 ETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLEGNASVLAGPTGAGKTHMLGSPQPG 119  
Db 62 KDLKFFVDAVDFDETSTQSEVFHTTKPILRSFLNGYNTVLAAYGATGAGKTHMLGSADE 121



```
QY 120 PGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDPASGDLVIREDCRG 179
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 PGVWYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL-VNSGPLAVREDTQK 177

QY 180 NILIPGLS-OKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLA 237
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GVVVHGLTLHQPSS--EELHLLDNGNKNRTQHPDTMNATSSRSRSHAVFQIYLRQDKTA 235

QY 238 PFRQ--REGKLYLIDLAGEEDNRRTGNKGLRKESCAINTSLFVLGKVVDAL----- 287
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 SINQNVRIAKMSLIDLAGEASTSGAKGTRFVEGTNINRSLLALGNVINALADSKISFL 295

QY 288 -----NQGLPRVPYRDSKLTRELQDSLGSAHSILIANIAPERFFYLDTVSALNFAARSK 342
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 CFQKQKQ---HIPYRNSKLTRELKDSLGNCQTIMIAAVSPSSVFYDDTVNTLYKANRAK 352

QY 343 EV 344
||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 DI 354

RESULT 11
US-09-723-097-4
; Sequence 4, Application US/09723097
; Patent No. 6492151
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6492151el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/723,097
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Human
US-09-723-097-4

Query Match 31.1%; Score 549; DB 4; Length 355;
Best Local Similarity 39.0%; Pred. No. 4.6e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDGTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 MKVVRVRPENTKEKAAGFHKKVHVVDKHLVDFDPKQEEVSFPFHGKTTNQNVIKQN-- 61

QY 60 ETLKYQFDAPYGERSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPQ 119
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KDLKFVDAVDFETSTQSEVFEHTTAPILRSPLNGYNCVTLAYGATGAGKTHMLGSADE 121

QY 120 PGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDPASGDLVIREDCRG 179
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 PGVWYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL-VNSGPLAVREDTQK 177

QY 180 NILIPGLS-OKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLA 237
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GVVVHGLTLHQPSS--EELHLLDNGNKNRTQHPDTMNATSSRSRSHAVFQIYLRQDKTA 235

QY 238 PFRQ--REGKLYLIDLAGEEDNRRTGNKGLRKESCAINTSLFVLGKVVDAL----- 287
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 SINQNVRIAKMSLIDLAGEASTSGAKGTRFVEGTNINRSLLALGNVINALADSKISFL 295

QY 288 -----NQGLPRVPYRDSKLTRELQDSLGSAHSILIANIAPERFFYLDTVSALNFAARSK 342
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 CFQKQKQ---HIPYRNSKLTRELKDSLGNCQTIMIAAVSPSSVFYDDTVNTLYKANRAK 352

QY 343 EV 344
||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 DI 354

RESULT 13
US-09-724-511-2
; Sequence 2, Application US/09724511
; Patent No. 6391601
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391601el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/724,511
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
Db 353 DI 354

RESULT 12
US-09-632-344-4
; Sequence 4, Application US/09632344
; Patent No. 6534309
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6534309el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/632,344
; CURRENT FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Human
US-09-632-344-4

Query Match 31.1%; Score 549; DB 4; Length 355;
Best Local Similarity 39.0%; Pred. No. 4.6e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDGTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 MKVVRVRPENTKEKAAGFHKKVHVVDKHLVDFDPKQEEVSFPFHGKTTNQNVIKQN-- 61

QY 60 ETLKYQFDAPYGERSTQODIYAGSVOPILRHLEGGONASVLAYGPTGAGKTHMLGSPQ 119
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KDLKFVDAVDFETSTQSEVFEHTTAPILRSPLNGYNCVTLAYGATGAGKTHMLGSADE 121

QY 120 PGVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDPASGDLVIREDCRG 179
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 PGVWYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL-VNSGPLAVREDTQK 177

QY 180 NILIPGLS-OKPISSPADFERHFLP-ASRNRVTGATRLNQRSSRSRSHAVLLVKVDQERLA 237
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 GVVVHGLTLHQPSS--EELHLLDNGNKNRTQHPDTMNATSSRSRSHAVFQIYLRQDKTA 235

QY 238 PFRQ--REGKLYLIDLAGEEDNRRTGNKGLRKESCAINTSLFVLGKVVDAL----- 287
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 SINQNVRIAKMSLIDLAGEASTSGAKGTRFVEGTNINRSLLALGNVINALADSKISFL 295

QY 288 -----NQGLPRVPYRDSKLTRELQDSLGSAHSILIANIAPERFFYLDTVSALNFAARSK 342
|||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 CFQKQKQ---HIPYRNSKLTRELKDSLGNCQTIMIAAVSPSSVFYDDTVNTLYKANRAK 352

QY 343 EV 344
||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 353 DI 354

RESULT 13
US-09-724-511-2
; Sequence 2, Application US/09724511
; Patent No. 6391601
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391601el motor proteins and methods for
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/724,511
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Human
US-09-724-511-2

Query Match
  31.1%; Score 549; DB 3; Length 367;
Best Local Similarity 39.0%; Pred. No. 4.9e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWNRHQ 59
Db 12 MKVVVRVPENTKEKAAGFKVHVVDKHLVDFDPKQEEVSFFHGKTTNQNVIKQN-- 69
QY 60 ETLKYQDFAFGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPEQ 119
Db 70 KDLKFVDFDAVDETSTQSEVFEHTTKPIRLSFLNGYNTVLAYGATGAGKTHMLGSADE 129
QY 120 PGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYOEKVLDDLDPASGDLVIREDCRG 179
Db 130 PGVMYLTMLHLYKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL--VNSGFLAVREDTQK 185
QY 180 NILIPGLS--QKPISSFADPERHFLP--ASRNTVGTATRLNQSRSSSHAVLLVKVDQERLA 237
Db 186 GVVVHGLTLHQPSS--EEILHLLDNGNKNRTOHPTDMNATSSRSHAVFQIYLRQODKTA 243
QY 238 PFRQ--REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDAL----- 287
Db 244 SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLLAGNVINALADSKISFL 303
QY 288 -----NOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSK 342
Db 304 CFORKNQ---HIPYRNSKLTLLKDSLGNCQOTIMAAVSPSSVFYDDTYNTLKYANRAK 360
QY 343 EV 344
Db 361 DI 362

RESULT 15
US-09-632-344-2
; Sequence 2, Application US/09632344
; Patent No. 6534309
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6534309el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/632,344
; CURRENT FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Human
US-09-632-344-2

Query Match
  31.1%; Score 549; DB 4; Length 367;
Best Local Similarity 39.0%; Pred. No. 4.9e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWNRHQ 59
Db 12 MKVVVRVPENTKEKAAGFKVHVVDKHLVDFDPKQEEVSFFHGKTTNQNVIKQN-- 69
QY 60 ETLKYQDFAFGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPEQ 119
Db 70 KDLKFVDFDAVDETSTQSEVFEHTTKPIRLSFLNGYNTVLAYGATGAGKTHMLGSADE 129
QY 120 PGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYOEKVLDDLDPASGDLVIREDCRG 179
Db 130 PGVMYLTMLHLYKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL--VNSGFLAVREDTQK 185
QY 180 NILIPGLS--QKPISSFADPERHFLP--ASRNTVGTATRLNQSRSSSHAVLLVKVDQERLA 237
Db 186 GVVVHGLTLHQPSS--EEILHLLDNGNKNRTOHPTDMNATSSRSHAVFQIYLRQODKTA 243
QY 238 PFRQ--REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVVDAL----- 287
Db 244 SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLLAGNVINALADSKISFL 303
QY 288 -----NOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSK 342
Db 304 CFORKNQ---HIPYRNSKLTLLKDSLGNCQOTIMAAVSPSSVFYDDTYNTLKYANRAK 360
QY 343 EV 344
Db 361 DI 362

RESULT 14
US-09-723-097-2
; Sequence 2, Application US/09723097
; Patent No. 6492151
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6492151el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1046
; CURRENT APPLICATION NUMBER: US/09/723,097
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/632,344
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Human
US-09-723-097-2

Query Match
  31.1%; Score 549; DB 4; Length 367;
Best Local Similarity 39.0%; Pred. No. 4.9e-50;
Matches 141; Conservative 59; Mismatches 114; Indels 48; Gaps 11;

QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWNRHQ 59
Db 12 MKVVVRVPENTKEKAAGFKVHVVDKHLVDFDPKQEEVSFFHGKTTNQNVIKQN-- 69
QY 60 ETLKYQDFAFGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPEQ 119
Db 70 KDLKFVDFDAVDETSTQSEVFEHTTKPIRLSFLNGYNTVLAYGATGAGKTHMLGSADE 129
```

Search completed: November 5, 2004, 18:46:46  
Job time : 14.5172 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 19.0257 Seconds  
(without alignments)  
1697.543 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPPARV.....LEAKMLAKAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCFUS COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	3	US-09-724-224-8
2	2472	100.0	487	4	US-10-093-317-8
3	2467	99.8	512	3	US-09-724-224-4
4	2467	99.8	512	4	US-10-093-317-4
5	2467	99.8	665	4	US-09-595-684B-35
6	1768	71.5	346	3	US-09-724-224-6
7	1768	71.5	346	4	US-10-093-317-6
8	1763	71.3	370	3	US-09-724-224-2
9	1763	71.3	370	4	US-10-093-317-2
10	573.5	23.2	1234	4	US-09-592-054-8
11	572	23.1	522	4	US-09-592-054-4
12	568.5	23.0	473	4	US-09-592-054-6
13	561.5	22.7	1232	4	US-09-592-054-2
14	561	22.7	1388	3	US-09-572-191-2
15	561	22.7	1388	3	US-09-723-262-2
16	556.5	22.7	1388	3	US-09-723-219-2
17	556.5	22.5	1066	3	US-09-541-782-8
18	556.5	22.5	1066	4	US-09-723-820-8
19	556.5	22.5	1066	4	US-10-270-085-8
20	554	22.4	1279	3	US-09-724-517-2
21	554	22.4	1279	4	US-09-641-807A-2
22	554	22.4	1279	4	US-09-723-096-2
23	553.5	22.4	1231	4	US-09-595-684B-23
24	551	22.3	2954	4	US-09-150-867-1
25	550	22.2	864	4	US-09-883-096-2
26	549	22.2	355	3	US-09-724-511-4
27	549	22.2	355	4	US-09-723-097-4

28	549	22.2	355	4	US-09-632-344-4	Sequence 4, Appli
29	549	22.2	367	3	US-09-724-511-2	Sequence 2, Appli
30	549	22.2	367	4	US-09-723-097-2	Sequence 2, Appli
31	549	22.2	367	4	US-09-632-344-2	Sequence 2, Appli
32	546.5	22.1	1103	3	US-09-162-373-1	Sequence 1, Appli
33	546.5	22.1	1103	3	US-09-467-946-1	Sequence 1, Appli
34	544.5	22.0	935	4	US-09-914-259-25	Sequence 25, Appli
35	543	22.0	409	3	US-09-572-191-6	Sequence 6, Appli
36	543	22.0	409	3	US-09-723-262-6	Sequence 6, Appli
37	543	22.0	409	3	US-09-723-219-6	Sequence 6, Appli
38	543	22.0	928	4	US-09-914-259-23	Sequence 23, Appli
39	541	21.9	1690	4	US-09-595-684B-39	Sequence 39, Appli
40	535	21.6	513	4	US-09-724-519-6	Sequence 6, Appli
41	535	21.6	513	4	US-09-592-037-6	Sequence 6, Appli
42	535	21.6	513	4	US-09-428-156B-6	Sequence 2, Appli
43	533.5	21.6	1057	4	US-09-428-156B-2	Sequence 29, Appli
44	532	21.5	1056	4	US-09-595-684B-29	Sequence 10, Appli
45	532	21.5	1057	3	US-09-541-782-10	

ALIGNMENTS

RESULT 1  
US-09-724-224-8  
; Sequence 8, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724,224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-224-8

Query Match	100.0%	Score	2472;	DB	3;	Length	487;
Best Local Similarity	100.0%	Pred. No.	1.1e-225;				
Matches	487;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MGRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60				
Db	1	MGRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60				
Qy	61	TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLBEGQNASVLAYGPTGAGKTHMLGSPEQP	120				
Db	61	TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLBEGQNASVLAYGPTGAGKTHMLGSPEQP	120				
Qy	121	GVTPRALMDLLQITRECAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVITREDCRGN	180				
Db	121	GVTPRALMDLLQITRECAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVITREDCRGN	180				
Qy	181	ILIPGLSOKPISGFADPFERHFLPASNRRTVGATRLNORSSRSHAVLVKVDQERLAPFR	240				
Db	181	ILIPGLSOKPISGFADPFERHFLPASNRRTVGATRLNORSSRSHAVLVKVDQERLAPFR	240				
Qy	241	QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGVKVDALNQGPRVPRYRDSK	300				
Db	241	QREGKLYLIDLAGSEDNRRRTGNKGLRKESGAINTSLFVLGVKVDALNQGPRVPRYRDSK	300				
Qy	301	LTRLLODSLGSSAHSILIANIAPERRPYLDTVSALNFAARSKVINRPFPTNESLQPHALG	360				
Db	301	LTRLLODSLGSSAHSILIANIAPERRPYLDTVSALNFAARSKVINRPFPTNESLQPHALG	360				
Qy	361	PVKLSQKELGLGPEAKRGPPEEETGSPFMAAPASASQKLSPLQKLSMDPAMLERLL	420				

Db 361 PVKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 420  
QY 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVBEKDLIERLTKTKQKELEAKMLAQAEK 480  
Db 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVBEKDLIERLTKTKQKELEAKMLAQAEK 480  
QY 481 ENHCPTM 487  
Db 481 ENHCPTM 487

RESULT 2  
US-10-093-317-8  
; Sequence 8, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; CURRENT FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Human  
US-10-093-317-8

Query Match 100.0%; Score 2472; DB 4; Length 487;  
Best Local Similarity 100.0%; Pred. No. 1.1e-225;  
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
Db 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
QY 61 TLKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSP 120  
Db 61 TLKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSP 120  
QY 121 GVTPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDC 180  
Db 121 GVTPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDC 180  
QY 181 ILTPGLSOKPISSFADFERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAP 240  
Db 181 ILTPGLSOKPISSFADFERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAP 240  
QY 241 QREGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDS 300  
Db 241 QREGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDS 300  
QY 301 LTRLLQDSLGSSAHSILIANIAPERFYLDTSALNFAARSKVINRPTNESLOPHALG 360  
Db 301 LTRLLQDSLGSSAHSILIANIAPERFYLDTSALNFAARSKVINRPTNESLOPHALG 360  
QY 361 PVKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 420  
Db 361 PVKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 420  
QY 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVBEKDLIERLTKTKQKELEAKMLAQAEK 480  
Db 421 SLDRLLASQSGCAPLLSTPKRRMVLMTVBEKDLIERLTKTKQKELEAKMLAQAEK 480  
QY 481 ENHCPTM 487  
Db 481 ENHCPTM 487

RESULT 3  
US-09-724-224-4  
; Sequence 4, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724,224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-224-4

Query Match 99.8%; Score 2467; DB 3; Length 512;  
Best Local Similarity 100.0%; Pred. No. 3.7e-225;  
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GRCRLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 61  
Db 27 GRCRLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 86  
QY 62 LKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSP 121  
Db 87 LKYQDAFYGERSTQODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSP 146  
QY 122 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDC 181  
Db 147 VIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDC 206  
QY 182 ILTPGLSOKPISSFADFERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAP 241  
Db 207 ILTPGLSOKPISSFADFERHFLPASNRRTVGATRLNQRSRSHAVLLVKVDQERLAP 266  
QY 242 REGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDS 301  
Db 267 REGKLYLIDLAGSENRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPYRDS 326  
QY 302 TRLLODSLGSSAHSILIANIAPERFYLDTSALNFAARSKVINRPTNESLOPHALG 361  
Db 327 TRLLODSLGSSAHSILIANIAPERFYLDTSALNFAARSKVINRPTNESLOPHALG 386  
QY 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 421  
Db 387 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLL 446  
QY 422 LDRLLASQSGCAPLLSTPKRRMVLMTVBEKDLIERLTKQKELEAKMLAQAEK 481  
Db 447 LDRLLASQSGCAPLLSTPKRRMVLMTVBEKDLIERLTKQKELEAKMLAQAEK 506  
QY 482 NHCPTM 487  
Db 507 NHCPTM 512

RESULT 4  
US-10-093-317-4  
; Sequence 4, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317

```

; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human
; US-10-093-317-4

Query Match
Best Local Similarity 99.8%; Score 2467; DB 4; Length 512;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQT 61
Db 27 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQT 86
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMLGSPQPG 121
Db 87 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMLGSPQPG 146
QY 122 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
Db 147 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206
QY 182 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLLVKVDORERLAPRQ 241
Db 207 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLLVKVDORERLAPRQ 266
QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301
Db 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 326
QY 302 TRLLQDSLGSASIIILANIPARRFYLDTVSALNFAARKEVINRPTNESLQPHALGP 361
Db 327 TRLLQDSLGSASIIILANIPARRFYLDTVSALNFAARKEVINRPTNESLQPHALGP 386
QY 362 VKLSQKELGPPPEAKRGPEEEIGSPPEMAAPASASOKLSPLOKLSMDPAMLERLLS 421
Db 387 VKLSQKELGPPPEAKRGPEEEIGSPPEMAAPASASOKLSPLOKLSMDPAMLERLLS 446
QY 422 LDRLLASQSGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKAEKE 481
Db 447 LDRLLASQSGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKAEKE 506
QY 482 NHCPTM 487
Db 507 NHCPTM 512

```

```

RESULT 5
US-09-595-684B-35
; Sequence 35, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35

```

```

; LENGTH: 665
; TYPE: PRT
; ORGANISM: Human
; US-09-595-684B-35

Query Match
Best Local Similarity 99.8%; Score 2467; DB 4; Length 665;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQT 61
Db 26 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQT 85
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMLGSPQPG 121
Db 86 LKQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAAYGPTGAGKTHMLGSPQPG 145
QY 122 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181
Db 146 VIPALMDLLQLTREAGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 205
QY 182 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLLVKVDORERLAPRQ 241
Db 206 LIPCLSKQPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLLVKVDORERLAPRQ 265
QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 301
Db 266 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGLPRVPYRDSKL 325
QY 302 TRLLQDSLGSASIIILANIPARRFYLDTVSALNFAARKEVINRPTNESLQPHALGP 361
Db 326 TRLLQDSLGSASIIILANIPARRFYLDTVSALNFAARKEVINRPTNESLQPHALGP 385
QY 362 VKLSQKELGPPPEAKRGPEEEIGSPPEMAAPASASOKLSPLOKLSMDPAMLERLLS 421
Db 386 VKLSQKELGPPPEAKRGPEEEIGSPPEMAAPASASOKLSPLOKLSMDPAMLERLLS 445
QY 422 LDRLLASQSGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKAEKE 481
Db 446 LDRLLASQSGAPLLSTPKRERVMKMTVEEKDLEIRLTKQKELEAKMLAQKAEKE 505
QY 482 NHCPTM 487
Db 506 NHCPTM 511

```

```

RESULT 6
US-09-724-224-6
; Sequence 6, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: their use
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Human
; US-09-724-224-6

Query Match
Best Local Similarity 71.5%; Score 1768; DB 3; Length 346;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANRNHQT 60

```

Db 1 MGRCLSKIGATERPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
QY 61 TLKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQOP 120  
Db 61 TLKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQOP 120  
QY 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
Db 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
QY 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240  
Db 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240  
QY 241 QREGKLYLIDLAGESEDNRRRTGNKGLKESGAINTSLFVLGKVVDAALNOGLPRVPYRDSK 300  
Db 241 QREGKLYLIDLAGESEDNRRRTGNKGLKESGAINTSLFVLGKVVDAALNOGLPRVPYRDSK 300  
QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
Db 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

## RESULT 7

US-10-093-317-6  
; Sequence 6, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; CURRENT FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Human  
US-10-093-317-6

Query Match 71.5%; Score 1768; DB 4; Length 346;  
Best Local Similarity 100.0%; Pred. No. 4e-159;  
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATERPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
Db 1 MGRCLSKIGATERPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
QY 61 TLKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQOP 120  
Db 61 TLKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQOP 120  
QY 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
Db 121 GVTPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
QY 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240  
Db 181 ILIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 240  
QY 241 QREGKLYLIDLAGESEDNRRRTGNKGLKESGAINTSLFVLGKVVDAALNOGLPRVPYRDSK 300  
Db 241 QREGKLYLIDLAGESEDNRRRTGNKGLKESGAINTSLFVLGKVVDAALNOGLPRVPYRDSK 300  
QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
Db 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346

## RESULT 8

US-09-724-224-2  
; Sequence 2, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724,224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-224-2

Query Match 71.3%; Score 1763; DB 3; Length 370;  
Best Local Similarity 100.0%; Pred. No. 1.3e-158;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATERPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHOE 61  
Db 26 GRCRLSKIGATERPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHOE 85  
QY 62 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQOP 121  
Db 86 LKYQDFAFYGRSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSPQOP 145  
QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181  
Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205  
QY 182 LIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 241  
Db 206 LIPGLSQPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPFR 265  
QY 242 REGKLYLIDLAGESEDNRRRTGNKGLKESGAINTSLFVLGKVVDAALNOGLPRVPYRDSK 301  
Db 266 REGKLYLIDLAGESEDNRRRTGNKGLKESGAINTSLFVLGKVVDAALNOGLPRVPYRDSK 325  
QY 302 TRLLODSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
Db 326 TRLLODSLGSSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370

## RESULT 9

US-10-093-317-2  
; Sequence 2, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; CURRENT FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 370  
; TYPE: PRT  
; ORGANISM: Human  
US-10-093-317-2

Query Match 71.3%; Score 1763; DB 4; Length 370;  
Best Local Similarity 100.0%; Pred. No. 1.3e-158;



Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPPPRRVAVRLRPVDTAGASDPPCVRGMDSCSLEIANWRNHOET 61  
 Db 26 GRCLSKIGATRRPPPPRRVAVRLRPVDTAGASDPPCVRGMDSCSLEIANWRNHOET 85  
 Qy 62 LKYPDAPYGRSTQDDIYAGSVQPIILHLLLEGONASVLAIGPTGAGKTHMLGSPQPG 121  
 Db 86 LKYPDAPYGRSTQDDIYAGSVQPIILHLLLEGONASVLAIGPTGAGKTHMLGSPQPG 145  
 Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181  
 Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205  
 Qy 182 LIPGLSQKPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPRQ 241  
 Db 206 LIPGLSQKPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPRQ 265  
 Qy 242 REGKYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGRLPRVPRYRDSKL 301  
 Db 266 REGKYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGRLPRVPRYRDSKL 325  
 Qy 302 TRLLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
 Db 326 TRLLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370

RESULT 10  
 US-09-592-054-8  
 ; Sequence 8, Application US/09592054  
 ; Patent No. 6440684  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Finer, Jeffrey  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Wood, Kenneth  
 ; TITLE OF INVENTION: No. 6440684el motor proteins and methods for  
 ; FILE REFERENCE: 1016  
 ; CURRENT APPLICATION NUMBER: US/09/592,054  
 ; CURRENT FILING DATE: 2000-07-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 1234  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-592-054-8

Query Match 23.2%; Score 573.5; DB 4; Length 1234;  
 Best Local Similarity 32.2%; Pred. No. 6.1e-45;  
 Matches 156; Conservative 97; Mismatches 173; Indels 59; Gaps 13;

Qy 20 VRVAVRLRPVDTAGASDPPC-----VRGMDSCSLEIANWRNHOETLKYQDAF 69  
 Db 10 VRVALRCRPLVPKEISGCQCLSFVPGTQVVVGTDK-----SFTYDFV 54  
 Qy 70 YGERSTQDDIYAGSVQPIILHLLLEGONASVLAIGPTGAGKTHMLGS-----PEQP--GV 122  
 Db 55 FDPCTEQEVRFNKAVAPLIGIFKGYNATVLAIGPTGAGKTHMLGS-----PEQP--GV 114  
 Qy 123 IPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 180  
 Db 115 IPR-----VIQLLKEIDQKDFEFTLVKSVLEIYNEIILDLCPREKAQINIREDPK 170  
 Qy 181 ILIPGLSQKPISSFADFERRHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPR 240  
 Db 171 IKIVGLTEKTVLVALDVTSCLEQGNRSRTVASTAMNSQSSRSRSHAIPTISLEQCKSKNS 230  
 Qy 241 QREGKYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKVDALNQGRLPR--VPYRD 298  
 Db 231 SFRSKHLVLAGSERQKTKAEGDRUKGININRGLLCLGNVISALGDDKKGKGFVPRD 290

Qy 299 SKLTRLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNPTNBSLOPHA 358  
 Db 291 SKLTRLQDSIGGSHLTMIACTVSPADSNLEETLSTLYADRAKKNKPIVN--IDPHT 348  
 Qy 359 LGPVLSQKELLGPEAKRARGPBEIEIGSPPEMAAPASASQKLSPTQKLSMDPAMLER 418  
 Db 349 AELNHLKQ-----QVQQLQVLLLOAHGGTLPGSINAEPSEN---LQSLMEKQNSLVEE 398  
 Qy 419 LLSLDRLLASQSGGAPLLSTPKREVMVLMKTVEEK--DLETERLKTQK--ELEAKMLAQK 476  
 Db 399 NEKLSRCLSKAAGTAQML-----ERILLTEQVNEKLNKLEELRQHAACKLDLQKJVT 453  
 Qy 477 AEEKE 481  
 Db 454 LEDQE 458

RESULT 11  
 US-09-592-054-4  
 ; Sequence 4, Application US/09592054  
 ; Patent No. 6440684  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Finer, Jeffrey  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Wood, Kenneth  
 ; TITLE OF INVENTION: No. 6440684el motor proteins and methods for  
 ; FILE REFERENCE: 1016  
 ; CURRENT APPLICATION NUMBER: US/09/592,054  
 ; CURRENT FILING DATE: 2000-07-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 522  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-592-054-4

Query Match 23.1%; Score 572; DB 4; Length 522;  
 Best Local Similarity 31.8%; Pred. No. 2.1e-45;  
 Matches 165; Conservative 99; Mismatches 167; Indels 88; Gaps 17;

Qy 1 MGRCLSKIGATRRPPPA-----RVRAVRLRP-----FVDTAGASD 38  
 Db 10 MGRIRI-----RAPSTSLSEEVGIPVRVALRCRPLVPKEISECCQCLSFVPG-----E 58  
 Qy 39 PPCVRGMDCSLEIANWRNHOETLKYQDAFYGERSTQDDIYAGSVQPIILHLLLEGONAS 98  
 Db 59 PQVVVGTDK-----SFTYDFVDPSTEQEVEFNTAVAPLIGKVFKNAT 103  
 Qy 99 VLAIGPTGAGKTHMLGS-----PEQP--GVIPRALMDLLQLTREEGAGRPWALSVTMS 151  
 Db 104 VLAIGPTGAGKTHMLGS-----PEQP--GVIPRALMDLLQLTREEGAGRPWALSVTMS 159  
 Qy 152 YLEIYQEKVLDLDPASGDLVIREDCRNIIPGLSQKPISSFADFERRHFLPASNRRT 209  
 Db 160 YLEIYNEIILDLCPREKAQINIREDPKEGKIVGLTEKTVLVALDVTSCLEQGNRSRT 219  
 Qy 210 VGATRLNQRSSRSRSHAVLLVKVDQERLAPPRQREGKYLIDLAGSDNRRTGNKGLRKE 269  
 Db 220 VASTAMNSQSSRSRSHAIPTISLEQCKSKNSFRSKLHLVDLAGSERQKTKAEGDRUK 279  
 Qy 270 SGAINSTSLFVLGKVDALNQGRLPR--VPYRDSKLTRLQDSIGGSAHSILIANIAPERRP 327  
 Db 280 GININRGLLCLGNVISALGDDKKGKGFVPRYRDSKLTRLQDSIGGSHLTMIACTVSPADSN 339  
 Qy 328 YLDTVSAALNFAARSKEVINRPTNPTNBSLOPHALGPVK--LSQKELGPEAKRARGPBEIE 385  
 Db 340 LESTLTLNRYADRAKKNKPIVNIDPQTALNHLKQVQVQLV-----LLQAHGG----- 391  
 Qy 386 IGSFPEMAAPASASQKLSL--LQKLSMDPAMLERLISLDRLLASQSGGAPLSTPKRER 444

```

392 -----TLPGSITVPESENLSQSLMEKNSQSLVEENKLSRGLSEAAQTQAML-----ER 439
445 MVMKTVBEK-DLEIERLTKQK-ELEAKMLAQAEKE 481
440 IITWEQANEKNAKLEELRQHAACKLDLQKLVETLEDOE 478

RESULT 12
US-09-592-054-6
; Sequence 6, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-6

Query Match 23.0%; Score 568.5; DB 4; Length 473;
Best Local Similarity 32.3%; Pred. No. 3.9e-45;
Matches 159; Conservative 97; Mismatches 162; Indels 75; Gaps 15;

QY 20 VRVAVRLRP-----FVDGTAGASDPPCVRGMDCSLEIANWRNHQETLKY 64
DB 8 VRVALRCRLPVKEISEGCMCLSFVPG-----EPQVVVGTDK-----SF 47
QY 65 QFDIFYGERSTQDDIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGS-----PEQ 119
DB 48 TYDFVDFPSTQEVEVNTAVAPLKGFKGNATVLAAYGQTSKTYSMGGAYTABQENE 107
QY 120 P--GVIPRALMDLLQTLRECAEGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
DB 108 PTVGVIPIR-----VIQLLFEIDKSDFEFTLKVSYLEIYNEEILDLLCPSREKAQINIRE 163
QY 176 DCRGNILPGLSQKPISSFADFHERFLPASRNTVGATRLNQRSSRSHAVLLVKVDQER 235
DB 164 DPKEGKIVGLTEKTVLVALDTVSCLEQGNNSRTVASTAMNSQSSRSHAIPTISLEQKK 223
QY 236 LAPFRQREGKLYLDLAGESENRRTGNKGLRKEGSAINTSLFVLGKVVVDALNQGLPR-- 293
DB 224 SDKNSFRSKLHLVDLAGSERQKTKAEGDRKKEGININRGLLCLGNVISALGDDKKGPF 283
QY 294 VPYRDSKLTLLQDSLGSAHSILIANIAPERRFYDVTVSALNFAARKEVINRPFNTES 353
DB 284 VPYRDSKLTLLQDSLGNSHTLMACVSPADSNLEETLNTLRVADRARKIKNKPINVID 343
QY 354 LQPHALGPVK--LSQKELLGPPEAKRARGPEEEIGSPPEMAAPASASQKLSL--LQKLSS 410
DB 344 PQTAELNHLKQVQQLQVL-----LLOAHGG-----TLPGSITVPESENLSQSLME 389
QY 411 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKQK-EL 468
DB 389 KNSQSLVEENKLSRGLSEAAQTQAML-----ERLIITWEQANEKNAKLEELRQHAACKL 443
QY 469 EAKMLAQAEKE 481
DB 444 DLQKLVEITLEDQE 456

RESULT 13
US-09-592-054-2
; Sequence 2, Application US/09592054
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17

```

```

; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Human
US-09-592-054-2

Query Match 22.7%; Score 561.5; DB 4; Length 1232;
Best Local Similarity 32.0%; Pred. No. 8.3e-44;
Matches 158; Conservative 97; Mismatches 163; Indels 75; Gaps 15;

QY 20 VRVAVRLRP-----FVDGTAGASDPPCVRGMDCSLEIANWRNHQETLKY 64
DB 10 VRVALRCRLPVKEISEGCMCLSFVPG-----EPQVVVGTDK-----SF 49
QY 65 QFDIFYGERSTQDDIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGS-----PEQ 119
DB 50 TYDFVDFPSTQEVEVNTAVAPLKGFKGNATVLAAYGQTSKTYSMGGAYTABQENE 109
QY 120 P--GVIPRALMDLLQTLRECAEGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
DB 110 PTVGVIPIR-----VIQLLFEIDKSDFEFTLKVSYLEIYNEEILDLLCPSREKAQINIRE 165
QY 176 DCRGNILPGLSQKPISSFADFHERFLPASRNTVGATRLNQRSSRSHAVLLVKVDQER 235
DB 166 DPKEGKIVGLTEKTVLVALDTVSCLEQGNNSRTVASTAMNSQSSRSHAIPTISLEQKK 225
QY 236 LAPFRQREGKLYLDLAGESENRRTGNKGLRKEGSAINTSLFVLGKVVVDALNQGLPR-- 293
DB 226 SDKNSFRSKLHLVDLAGSERQKTKAEGDRKKEGININRGLLCLGNVISALGDDKKGPF 285
QY 294 VPYRDSKLTLLQDSLGSAHSILIANIAPERRFYDVTVSALNFAARKEVINRPFNTES 353
DB 286 APYRDSKLTLLQDSLGNSHTLMACVSPADSNLEETLNTLRVADRARKIKNKPINVID 345
QY 354 LQPHALGPVK--LSQKELLGPPEAKRARGPEEEIGSPPEMAAPASASQKLSL--LQKLSS 410
DB 346 PQTAELNHLKQVQQLQVL-----LLOAHGG-----TLPGSITVPESENLSQSLME 390
QY 411 MDPAMLERLLSLDRLLASQSGCAPLLSTPKRERVMVLMKTVEEK-DLEIERLTKQK-EL 468
DB 391 KNSQSLVEENKLSRGLSEAAQTQAML-----ERLIITWEQANEKNAKLEELRQHAACKL 445
QY 469 EAKMLAQAEKE 481
DB 446 DLQKLVEITLEDQE 458

RESULT 14
US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17

```

NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1388  
; TYPE: PRT  
; ORGANISM: Human  
US-09-572-191-2

Query Match 22.7%; Score 561; DB 3; Length 1388;

Best Local Similarity 30.9%; Pred. No. 1.1e-43;  
Matches 162; Conservative 97; Mismatches 192; Indels 74; Gaps 15;

```
Qy 20 VRVAVRLRPFDV--GTAGASPPCVRGMDSCSLAIANWRNHQETLKYQFDIFYGERSTQQ 77
Db 27 IKFVIRIRPPAERSGSADGEQNCCLSVLSSTSLRL--HSNPEPKTFTFDHVADVDTTQE 83
Qy 78 DIVAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPREP-----GVIPRA---L 127
Db 84 SVFATVAKSIVESCSMGNGTIFAYGQTGSKTFTMMGPSEDNFSNHLRGVIRPSFEYL 143
Qy 128 MDLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNILIPGLS 187
Db 144 FSLIDREKEKAGAGKSFCK--CSFIEIYNEQIYDLDSDASAGLYLREHIKKGVPVVGAV 201
Qy 188 QKPISSPADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQRELRAPF--RQREGKL 246
Db 202 EQVVTSAEAYQVLGGWRNRVASTSMNRESSRSHAVFTTIESMEKSNIEIVNRTSLL 261
Qy 247 YLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDAL----NQGLPRVPYRDSKLT 302
Db 262 NLVDLAGSERQKQTHAEGMRLKEAGNINRSLSCLGQVITLVVDVGNKGQRHVYRDSKLT 321
Qy 303 RLLQDSLGGSASHILIANIAPERFYLDTVSALNFAARSKVINRPFNTNESLQPHALGPV 362
Db 322 FLRLDSLGGNAKTALIANVHPGSRFCGETLSTLNFPAQRAKLIKNAVVNEDTQ-----GNV 377
Qy 363 KLSQ-----KELIG-----PPEAKRARGPEEE---EIGSPPEMAAPASASOKLSPLQ 406
Db 378 SOLQAEVKRLKEQLAELASGQTPEPSFLTRDKKNTNMEYFQEAMLFFKKEQEKKSLE 437
Qy 407 KLSMDPAML-----ERLLSLDRLLASQSGA-----PLLSTPKR 442
Db 438 KVTQLEDLTLLKKEKFIOSNMIVKFREDQIIRLEKL--HKESRGGFLEPEQDRLLSELN 495
Qy 443 ERMVLMKTVEKD-----LEIERLKTQKELEAKMLAQAEEKE 481
Db 496 EIQTLEQIEHHPRVAKYAMENHSRENRRLRLLELPEVKRAQEMD 540
```

## RESULT 15

US-09-723-262-2  
; Sequence 2, Application US/09723262  
; Patent No. 6379912  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Wood, Kenneth  
; TITLE OF INVENTION: No. 6379912el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1017  
; CURRENT APPLICATION NUMBER: US/09/723,262  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US 09/572,191  
; PRIOR FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1388  
; TYPE: PRT  
; ORGANISM: Human  
US-09-723-262-2

Query Match

22.7%; Score 561; DB 3; Length 1388;

Best Local Similarity 30.9%; Pred. No. 1.1e-43;  
Matches 162; Conservative 97; Mismatches 192; Indels 74; Gaps 15;

```
Qy 20 VRVAVRLRPFDV--GTAGASPPCVRGMDSCSLAIANWRNHQETLKYQFDIFYGERSTQQ 77
Db 27 IKFVIRIRPPAERSGSADGEQNCCLSVLSSTSLRL--HSNPEPKTFTFDHVADVDTTQE 83
Qy 78 DIVAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGSPREP-----GVIPRA---L 127
Db 84 SVFATVAKSIVESCSMGNGTIFAYGQTGSKTFTMMGPSEDNFSNHLRGVIRPSFEYL 143
Qy 128 MDLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGNILIPGLS 187
Db 144 FSLIDREKEKAGAGKSFCK--CSFIEIYNEQIYDLDSDASAGLYLREHIKKGVPVVGAV 201
Qy 188 QKPISSPADFERHFLPASRNRTVGATRLNQSRSSHAVLLVKVDQRELRAPF--RQREGKL 246
Db 202 EQVVTSAEAYQVLGGWRNRVASTSMNRESSRSHAVFTTIESMEKSNIEIVNRTSLL 261
Qy 247 YLIDLAGSEDNRRTGNKGLRKESGAINTSFLVLGKVVDAL----NQGLPRVPYRDSKLT 302
Db 262 NLVDLAGSERQKQTHAEGMRLKEAGNINRSLSCLGQVITLVVDVGNKGQRHVYRDSKLT 321
Qy 303 RLLQDSLGGSASHILIANIAPERFYLDTVSALNFAARSKVINRPFNTNESLQPHALGPV 362
Db 322 FLRLDSLGGNAKTALIANVHPGSRFCGETLSTLNFPAQRAKLIKNAVVNEDTQ-----GNV 377
Qy 363 KLSQ-----KELIG-----PPEAKRARGPEEE---EIGSPPEMAAPASASOKLSPLQ 406
Db 378 SOLQAEVKRLKEQLAELASGQTPEPSFLTRDKKNTNMEYFQEAMLFFKKEQEKKSLE 437
Qy 407 KLSMDPAML-----ERLLSLDRLLASQSGA-----PLLSTPKR 442
Db 438 KVTQLEDLTLLKKEKFIOSNMIVKFREDQIIRLEKL--HKESRGGFLEPEQDRLLSELN 495
Qy 443 ERMVLMKTVEKD-----LEIERLKTQKELEAKMLAQAEEKE 481
Db 496 EIQTLEQIEHHPRVAKYAMENHSRENRRLRLLELPEVKRAQEMD 540
```

Search completed: November 5, 2004, 18:46:47  
Job time : 20.0257 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 82.6338 Seconds  
(without alignments)  
2081.726 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKMLAQKAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	16	US-10-797-893-8
2	2467	99.8	512	16	US-10-797-893-4
3	2309	93.4	490	15	US-10-334-143-80
4	2141	86.6	460	9	US-09-925-300-1228
5	1768	71.5	346	16	US-10-797-893-6
6	1763	71.3	370	16	US-10-797-893-2
7	608	24.6	589	17	US-10-425-115-320209
8	601	24.3	377	16	US-10-437-963-159031
9	578.5	23.4	1237	15	US-10-334-143-33
10	574.5	23.2	1232	14	US-10-116-712-670
11	574.5	23.2	1232	16	US-10-408-765A-2153
12	574.5	23.2	1235	15	US-10-334-143-8
13	571	23.1	776	16	US-10-437-963-103943

14	571	23.1	783	17	US-10-425-115-332919	Sequence 332919,
15	571	23.1	854	15	US-10-425-114-59708	Sequence 59708, A
16	564.5	22.8	548	15	US-10-108-260A-2692	Sequence 2692, Ap
17	563.5	22.8	1001	17	US-10-425-115-231231	Sequence 231231,
18	561	22.7	1388	14	US-10-146-473-82	Sequence 82, Appl
19	561	22.7	1388	15	US-10-173-999-32	Sequence 32, Appl
20	561	22.7	1388	15	US-10-332-089-2	Sequence 2, Appli
21	561	22.7	1388	15	US-10-188-832-164	Sequence 164, App
22	560	22.7	992	17	US-10-739-930-7903	Sequence 7903, Ap
23	559.5	22.6	1401	15	US-10-287-226-142	Sequence 142, App
24	559	22.6	1030	15	US-10-425-114-62748	Sequence 62748, A
25	557.5	22.6	672	16	US-10-408-765A-1664	Sequence 1664, Ap
26	557.5	22.6	1232	14	US-10-116-712-664	Sequence 664, App
27	557.5	22.6	1232	14	US-10-116-712-669	Sequence 669, App
28	557	22.5	1324	15	US-10-287-226-314	Sequence 314, App
29	554	22.4	757	15	US-10-220-120-366	Sequence 366, App
30	554	22.4	757	16	US-10-363-829-316	Sequence 316, App
31	550	22.2	864	9	US-09-883-096-2	Sequence 2, Appli
32	546.5	22.1	1103	9	US-09-847-874A-1	Sequence 1, Appli
33	546.5	22.1	1103	14	US-10-458-162-1	Sequence 1, Appli
34	546.5	22.1	1826	16	US-10-408-765A-1477	Sequence 1477, Ap
35	545.5	22.1	834	15	US-10-287-226-306	Sequence 306, App
36	544.5	22.0	935	14	US-10-080-608A-25	Sequence 25, Appl
37	544.5	22.0	935	14	US-10-370-685-114	Sequence 114, App
38	543	22.0	409	15	US-10-332-089-6	Sequence 6, Appli
39	543	22.0	928	14	US-10-080-608A-23	Sequence 23, Appl
40	543	22.0	928	14	US-10-370-685-112	Sequence 112, App
41	541.5	21.9	1382	16	US-10-437-963-176714	Sequence 176714,
42	535	21.6	513	17	US-10-601-036-6	Sequence 6, Appli
43	533.5	21.6	1057	17	US-10-601-036-2	Sequence 2, Appli
44	532	21.5	1056	14	US-10-282-174-472	Sequence 472, App
45	532	21.5	1056	14	US-10-282-174-474	Sequence 474, App

## ALIGNMENTS

RESULT 1  
US-10-797-893-8  
; Sequence 8, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Human  
; US-10-797-893-8

Query Match 100.0%; Score 2472; DB 16; Length 487;  
Best Local Similarity 100.0%; Pred No. 9.3e-202;  
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGRCLSKIGATRRPPPARVVRVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60
Db	1	MGRCLSKIGATRRPPPARVVRVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRHQE	60
Qy	61	TLKYQDFAFGERSTOODIYAGSVQPIRLHLLRGONASVLAYGPTGAGKTHMLGSPQ	120
Db	61	TLKYQDFAFGERSTOODIYAGSVQPIRLHLLRGONASVLAYGPTGAGKTHMLGSPQ	120
Qy	121	GVIPRALMDLLQLTREAGRGPWALSVTWTSYLEIYQKVLDDLDPASGDLVIREDCRGN	180

Db 121 GVIPRALMDLQLTREAGGPRWALSVTMSYLEIQKVLDDLPASGDLVIREDCGN 180  
 QY 181 ILIPGLSQPISFADFPRHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240  
 Db 181 ILIPGLSQPISFADFPRHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240  
 QY 241 QREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 300  
 Db 241 QREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDVTYALNFAARSKEVINRPTNESLOPHALG 360  
 Db 301 LTRLLQDSLGSAHSILIANIAPERRFYLDVTYALNFAARSKEVINRPTNESLOPHALG 360  
 QY 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
 Db 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
 QY 421 SLDRLLASQSGCAPLLSTPKRERMVLMKTVEKDLERLTKKQELAKMLAKAEBK 480  
 Db 421 SLDRLLASQSGCAPLLSTPKRERMVLMKTVEKDLERLTKKQELAKMLAKAEBK 480  
 QY 481 ENHCPTM 487  
 Db 481 ENHCPTM 487

RESULT 2

US-10-797-893-4  
 ; Sequence 4, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 512  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-10-797-893-4

Query Match 99.8%; Score 2467; DB 16; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-201;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 Db 27 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86  
 QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMTLGSPPQPG 121  
 Db 87 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMTLGSPPQPG 146  
 QY 122 VIPRALMDLQLTREAGGPRWALSVTMSYLEIQKVLDDLPASGDLVIREDCGN 181  
 Db 147 VIPRALMDLQLTREAGGPRWALSVTMSYLEIQKVLDDLPASGDLVIREDCGN 206  
 QY 182 LIPGLSQPISFADFPRHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 241  
 Db 207 LIPGLSQPISFADFPRHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 266  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 301

Db 267 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 326  
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDVTYALNFAARSKEVINRPTNESLOPHALGP 361  
 Db 327 TRLLQDSLGSAHSILIANIAPERRFYLDVTYALNFAARSKEVINRPTNESLOPHALGP 386  
 QY 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
 Db 387 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446  
 QY 422 LDRLLASQSGCAPLLSTPKRERMVLMKTVEKDLERLTKKQELAKMLAKAEBK 481  
 Db 447 LDRLLASQSGCAPLLSTPKRERMVLMKTVEKDLERLTKKQELAKMLAKAEBK 506  
 QY 482 NHCPTM 487  
 Db 507 NHCPTM 512

RESULT 3

US-10-334-143-80  
 ; Sequence 80, Application US/10334143  
 ; Publication No. US20040009549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
 ; APPLICANT: SUDARSANAM, SUCHA  
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
 ; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
 ; FILE REFERENCE: 038602/1543  
 ; CURRENT APPLICATION NUMBER: US/10/334,143  
 ; CURRENT FILING DATE: 2002-12-31  
 ; PRIOR APPLICATION NUMBER: 60/343,169  
 ; PRIOR FILING DATE: 2001-12-31  
 ; NUMBER OF SEQ ID NOS: 207  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 80  
 ; LENGTH: 490  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-334-143-80

Query Match 93.4%; Score 2309; DB 15; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-188;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 Db 34 GRCRLSKIGATRRPPPARVRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93  
 QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMTLGSPEQPG 121  
 Db 94 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMTLGSPEQPG 153  
 QY 122 VIPRALMDLQLTREAGGPRWALSVTMSYLEIQKVLDDLPASGDLVIREDCGN 181  
 Db 154 VIPRALMDLQLTREAGGPRWALSVTMSYLEIQKVLDDLPASGDLVIREDCGN 213  
 QY 182 LIPGLSQPISFADFPRHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 241  
 Db 214 LIPGLSQPISFADFPRHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 273  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 301  
 Db 274 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSK 333  
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDVTYALNFAARSKEVINRPTNESLOPHALGP 361  
 Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDVTYALNFAARSKEVINRPTNESLOPHALGP 393  
 QY 362 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
 Db 394 VKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 453

QY 422 LDRLLASQSGAPILLSTPKRERMVLMKTVBEKOL 456  
 Db 454 LDRLLASQSGAPILLSTPKRERMVLMKTVBEKOL 488

## RESULT 4

US-09-925-300-1228  
 ; Sequence 1228, Application US/09925300  
 ; Publication No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben,  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1228  
 ; LENGTH: 460  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (75)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (147)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (435)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

Query Match 86.6%; Score 2141; DB 9; Length 460;  
 Best Local Similarity 95.9%; Pred. No. 1.4e-173;  
 Matches 422; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 Db 4 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 63

QY 62 LKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 121  
 Db 64 LKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 123

QY 122 VIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181  
 Db 124 VIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 183

QY 182 LIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 241  
 Db 184 LIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ 243

QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSKL 301  
 Db 244 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSKL 303

QY 302 TRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLQPHALGP 361  
 Db 304 TRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLQPHALGP 363

QY 362 VKLSQKELLGPPPEAKRARGPEEIEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
 Db 364 VKLSQKELLGPPPEAKRARGPEEIEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLQ 423

QY 422 LDRLLASQSGAPILLSTPK 441  
 Db 424 LGPSACLPGEPXGFSVEYPK 443

## RESULT 5

US-10-797-893-6  
 ; Sequence 6, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 346  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-10-797-893-6

Query Match 71.5%; Score 1768; DB 16; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-142;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQE 60  
 Db 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQE 60

QY 61 TLKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 120  
 Db 61 TLKYQDFAFGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPQPG 120

QY 121 GVIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
 Db 121 GVIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180

QY 181 ILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240  
 Db 181 ILIPGLSQKPISSPADFERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240

QY 241 QREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 300  
 Db 241 QREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 300

QY 301 LTRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 Db 301 LTRLLQDSLGSASHILIANIAPERRFYLDTVSALNFAARKEVIN 346

## RESULT 6

US-10-797-893-2  
 ; Sequence 2, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 370

[illegible]

RESULT 8  
 US-10-437-963-159031  
 ; Sequence 159031, Application US/10437963  
 ; Publication NO. US2004012343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 159031  
 ; LENGTH: 377  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(377)  
 ; OTHER INFORMATION: unsure at all Xaa locations

```
i OTHER INFORMATION: Clone ID: PAT_MRT4530_58448C.1.psp
US-10-437-963-159031

Query Match          24.3%   Score 601; DB 16; Length 377;
Best Local Similarity 36.7%; Pred No. 2.le=42;
Matches 151; Conservative 76; Mismatches 133; Indels 52; Gaps 12

QY      6 LSKIGATRRPPPRVVRVAVRLRPFVDGTAGASDPPCVRGWMDS-CSLEIATNRNHOETLK- 63
       :: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
Db      1 MATAAATQSQP--VVRVLRVRPHLPSEANSAEAPCGVLGLSGHPGGEVTVOLKDYTSRN 57

QY      64 --YQDAFYGRSTQQDIYAGSYQPIILRHLLSQNASVLAYGPTGAGTKHTMLGSPEQFG 121
       |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
Db      58 ECKYKLDAFFGESRVCSEIFQESVASVPGIFETNAIVFAYGATSGKTYTMOQTEDLPG 117

QY      122 VTIPALMDLLQLTREGAEGRPALSVYMSLYLSYYOEKVLDLLDPASGDPLVTREDCRGN 181
       :: :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



Db 118 LMXSTVLACTGT-----WC-SVEISYEVYMERCVYDLSLBPFAREIMVLDDKGNL 167  
Qy 182 LIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQORERLAPRQ 241  
Db 168 QKGLAWPVRSLEFHEIYISIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 220  
Qy 242 REGKYLII---DLAGSEDRRTGKGLRLKESGAINTSLFVLGKVDALNOGLPRVPYRD 298  
Db 221 VKGKLNLTSDLAGNEDNRTCNREGIRLQBSAKINOSLFSALNKNKEPRIPYRE 280  
Qy 299 SKLTRLQDSLGSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHA 358  
Db 281 SKLTRLQDSLGSGSHAVMIACLNPVE--YQEAHVHTVSLAARSHRVN-----HM 328  
Qy 359 LGPVKLSQKELLGPPEAK-----RARGPEE-----EIGSPPEMAAPASASQ 400  
Db 329 SSASKXKDKVDM---EAKLRAWLESKGTKSIQRMWDGLLSFNAIKTFLSMSQ 377  
RESULT 9  
US-10-334-143-33  
; Sequence 33, Application US/10334143  
; Publication No. US20040009549A1  
; GENERAL INFORMATION:  
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1543  
; CURRENT APPLICATION NUMBER: US/10/334,143  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: 60/343,169  
; PRIOR FILING DATE: 2001-12-31  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 1237  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-334-143-33

Query Match 23.4%; Score 578.5; DB 15; Length 1237;  
Best Local Similarity 32.4%; Pred. No. 1e-39;  
Matches 157; Conservative 97; Mismatches 172; Indels 59; Gaps 13;  
Qy 20 VRVAVLRPFVDGTAGASDPPC-----VRGMDSCSLEIANWRNHQETLKYQDAP 69  
Db 13 VRVALRCRPLVPKEISGCMCLSFVPGETQVVVGTDK-----SFTYDFV 57  
Qy 70 YGERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMTLGS-----PEQP--GV 122  
Db 58 FDPCTEQEEVFNKAVAPLIKGIFKGYNAVTLAYGOTSGKTYSGGAYTAQEENETVGI 117  
Qy 123 IPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIREDCRGN 180  
Db 118 IPR-----VIQLLFEIKDCKSDFEFTLKVSYLEIYNEEILDLCPREKAQINIREPK 173  
Qy 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQORERLAPR 240  
Db 174 IKIVGLTEKTVLVALDVTVSCLEQGNNSRTVASTAMNSQSSRSHAIFTTISLEQRKDKNC 233  
Qy 241 QREGKLYLIDLAGSEDRRTGKGLRLKESGAINTSLFVLGKVDALNOGLPR--VPYRD 298  
Db 234 SFRSKLHLVDLAGSERQKTKAEGDLRKEGININRGLLCGNVISALGDDDKSGSFVYRD 293  
Qy 299 SKLTRLQDSLGSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNESLOPHA 358  
Db 294 SKLTRLQDSLGSGSHAVMIACLNPVE--YQEAHVHTVSLAARSHRVN-----HM 328  
Qy 359 LGPVKLSQKELLGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSSMDPAMLER 418  
Db 352 AELNHLKQ-----QVQQLQVLLQAHGTTLPFSINAEFSEN---LQSLMEKNQSLVEE 401

Qy 419 LLSLDRLASQSGQAPLLSTPPKRRMVLMTKTVBEK--DLEIERLTKTKQ--ELEARMQAQK 476  
Db 402 NEXLSRCLSKAAGTAQML-----ERILITEQVNEKLNKLEELRQHVACKLDLOKLAVET 456  
Qy 477 ABEKE 481  
Db 457 LEDQE 461  
RESULT 10  
US-10-116-712-670  
; Sequence 670, Application US/10116712  
; Publication No. US20030194764A1  
; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Switzer, Ann  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.568  
; CURRENT APPLICATION NUMBER: US/10/116,712  
; CURRENT FILING DATE: 2002-04-07  
; NUMBER OF SEQ ID NOS: 670  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 670  
; LENGTH: 1232  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-712-670  
Query Match 23.2%; Score 574.5; DB 14; Length 1232;  
Best Local Similarity 32.5%; Pred. No. 2.2e-39;  
Matches 160; Conservative 97; Mismatches 161; Indels 75; Gaps 15;  
Qy 20 VRVAVLRP-----FVDGTAGASDPPCVRGMDSCSLEIANWRNHQETLKY 64  
Db 10 VRVALRCRPLVPKEISGCMCLSFVPG-----EPQVVVGTDK-----SF 49  
Qy 65 QFDAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMTLGS-----PEQ 119  
Db 50 TYDFVDFDPSTEQEEVFNTAVAPLIKGIFKGYNAVTLAYGOTSGKTYSGGAYTAQEEN 109  
Qy 120 P--GVIPRALMDLLQLTREBAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175  
Db 110 PTGVGVIPIR---VIQLLFEIKDCKSDFEFTLKVSYLEIYNEEILDLCPREKAQINIRE 165  
Qy 176 DCRGNLILPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQOR 235  
Db 166 DPKEGKIVGLTEKTVLVALDVTVSCLEQGNNSRTVASTAMNSQSSRSHAIFTTISLEQRK 225  
Qy 236 LAPFRQREGKLYLIDLAGSEDRRTGKGLRLKESGAINTSLFVLGKVDALNOGLPR-- 293  
Db 226 SDKNSRFRSKLHLVDLAGSERQKTKAEGDLRKEGININRGLLCGNVISALGDDDKGGF 285  
Qy 294 VPYRDSKLTRLQDSLGSGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNES 353  
Db 286 VPYRDSKLTRLQDSLGSGSHAVMIACLNPVE--YQEAHVHTVSLAARSHRVN-----HM 328  
Qy 354 LQPHALGPVK--LSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKSS 410  
Db 346 PQTAEINHLKQVQQLQVLL-----LQAHGG-----TLPGSITVFPSENLSLME 390  
Qy 411 MDPAELERLSDRLASQSGQAPLLSTPPKRRMVLMTKTVBEK--DLEIERLTKTKQ--EL 468  
Db 391 KQOSIVVEENEKLSRGLSEBAAGTAQML-----ERILITEQVNEKLNKLEELRQHVACKLDLOKLAVET 456  
Qy 469 EAKMLAQKABEKE 481  
Db 446 DLQKLVETLEDQE 458  
RESULT 11  
US-10-408-765A-2153  
; Sequence 2153, Application US/10408765A

Fri Nov 12 12:26:27 2004

```
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2153
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-2153

Query Match      23.2%; Score 574.5; DB 16; Length 1232;
Best Local Similarity 32.5%; Pred. No. 2.2e-39;
Matches 160; Conservative 97; Mismatches 161; Indels 75; Gaps 15;

QY      20 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLSIANWRNHQETLKY 64
Db      10 VRVALRCPLVPKEISEGCMCLSFVPG-----EPQVVVGTDK-----SF 49
QY      65 QFDIFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGS-----PEQ 119
Db      50 TYDFVDFPSTEQEEVFNTAVAPLIKGVFKGNATVLAYGQTGSGKTYSGMGAYTAQENE 109
QY      120 P--GVIPRALMDLLQTLREEGAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
Db      110 PTGVGVIPIR-----VIQLLFKEIDKKSDFEFTLKVSYLEIYNEEILDLLCPSREKAQINIRE 165
QY      176 DCRGNILPGLSQKPISSFADFERHFLPASRNTVGTATRLNQRSSSHAVLLVKVDQER 235
Db      166 DPKEGKIVGLTEKTVLVALDVTSCLEQGNNSRTVASTAMNSQSSRSHAFITISLEQRKK 225
QY      236 LAPPRQREGKYLIDLAGSEDRNRTGNKGLRKESGAINTSFLVLGVKVDALNOQLPR-- 293
Db      226 SDKNSFSRKLHLVDLAGSERQKTKAEGDRLEKGININRGLLCLGNVISALGDDKGGF 285
QY      294 VPYRDSKLTLLQDSLGSASHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNES 353
Db      286 VPYRDSKLTLLQDSLGSNSHTLMACVSPADSNLEETLNTLRYADRAKIKNKPINVID 345
QY      354 LQPHALGPVK--LSQKELLGPPEAKRARGPEEEIIGSPPEMAAPASASOKLSP-LQKLSS 410
Db      346 PQTAEHLNHLKQVOQLQVL-----LLOAHGG-----TLPGSITVPESENLOSLME 390
QY      411 MDPAMLERLLSLDRLLASQSGCAPLSTPKRERMVLMKTVEEK-DLEIERLKTQK-EL 468
Db      391 KNQSLIVENEKLISLSEAGQTAQML-----ERIIITEQANEKNNAKLELRQHAACKL 445
QY      469 EAKMLAQKAEKE 481
Db      446 DLQKLVTLEDOE 458
```

```
RESULT 12
US-10-334-143-8
; Sequence 8, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
```

```
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1235
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-334-143-8

Query Match      23.2%; Score 574.5; DB 15; Length 1235;
Best Local Similarity 32.5%; Pred. No. 2.2e-39;
Matches 160; Conservative 97; Mismatches 161; Indels 75; Gaps 15;

QY      20 VRVAVRLRP-----FVDGTAGASDPPCVRGMDSCSLSIANWRNHQETLKY 64
Db      13 VRVALRCPLVPKEISEGCMCLSFVPG-----EPQVVVGTDK-----SF 52
QY      65 QFDIFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAYGPTGAGKTHMLGS-----PEQ 119
Db      53 TYDFVDFPSTEQEEVFNTAVAPLIKGVFKGNATVLAYGQTGSGKTYSGMGAYTAQENE 112
QY      120 P--GVIPRALMDLLQTLREEGAGRPWALSVTMSYLEIYQEKVLDLDDPA--SGDLVIRE 175
Db      113 PTGVGVIPIR-----VIQLLFKEIDKKSDFEFTLKVSYLEIYNEEILDLLCPSREKAQINIRE 168
QY      176 DCRGNILPGLSQKPISSFADFERHFLPASRNTVGTATRLNQRSSSHAVLLVKVDQER 235
Db      169 DPKEGKIVGLTEKTVLVALDVTSCLEQGNNSRTVASTAMNSQSSRSHAFITISLEQRKK 228
QY      236 LAPPRQREGKYLIDLAGSEDRNRTGNKGLRKESGAINTSFLVLGVKVDALNOQLPR-- 293
Db      229 SDKNSFSRKLHLVDLAGSERQKTKAEGDRLEKGININRGLLCLGNVISALGDDKGGF 288
QY      294 VPYRDSKLTLLQDSLGSASHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNES 353
Db      289 VPYRDSKLTLLQDSLGSNSHTLMACVSPADSNLEETLNTLRYADRAKIKNKPINVID 348
QY      354 LQPHALGPVK--LSQKELLGPPEAKRARGPEEEIIGSPPEMAAPASASOKLSP-LQKLSS 410
Db      349 PQTAEHLNHLKQVOQLQVL-----LLOAHGG-----TLPGSITVPESENLOSLME 393
QY      411 MDPAMLERLLSLDRLLASQSGCAPLSTPKRERMVLMKTVEEK-DLEIERLKTQK-EL 468
Db      394 KNQSLIVENEKLISLSEAGQTAQML-----ERIIITEQANEKNNAKLELRQHAACKL 448
QY      469 EAKMLAQKAEKE 481
Db      449 DLQKLVTLEDOE 461
```

```
RESULT 13
US-10-437-963-103943
; Sequence 103943, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 103943
```

LENGTH: 776

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_101324C.1.pap

US-10-437-963-103943

Query Match 23.1%; Score 571; DB 16; Length 776;

Best Local Similarity 31.5%; Pred. No. 2.2e-39;

Matches 152; Conservative 99; Mismatches 194; Indels 38; Gaps 10;

QY 18 ARVAVAVLRPFVDGTAGASDPPCVGRGMDSCSLEIANWRNHQETLK-----YQDAF 69

DB 204 SRIMVFLRPMRSRKEKDGASRSCVKIVNKDDVLTPEFASSETDYLRLKRVGRHFCFDS 263

QY 70 YGERSTQDIIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGSPQPGVTPRALMD 129

DB 264 FPDITTAQEVYSTTSDLVVEGLQGRNGTVFCYATGAGKTYTMLGTWESPGVWVLAIKD 323

QY 130 LLQLTREGAERGPWALSVMYSLEYIOEKVLDLDPASGDLVIREDCRGNILIPGLSQK 189

DB 324 LFTKVRQRSHDGNH---SIQLSYLEVNETVRDLSLPCR-PLLRDKQGTV-AAGLTHY 378

QY 190 PISFADPERHFLPASNRRTVGATRLNORSRSHAVLLVKVDQRE-RLAPPRQEGKLYL 248

DB 379 RAYSTDVEMKLLQGNQNRTEPTRVNETSSRSHAILQVIVVEYSIDGSGIVTRVGKLSL 438

QY 249 IDLAGSDNRRTGKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRDSKLTLLQDS 308

DB 439 IDLAGSERALATDQRTQSIIEGANINRSLALSSCINALVEGKHIPYRNSKLTOLLKDS 498

QY 309 LGGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRNPNTNBSLOPHALGPVKLSQKE 368

DB 499 LGGSCNTVMIANISPSNLSFGETQNTLHWADRAKEIKTKALTANEVLRVTDSETQAK 558

QY 369 LL-----GPPRAKARPEEIEIGSPPEMAAPASAKLSPLQKLSNDPAMLERLLSLD 423

DB 559 LVLELOKENSELROQLARQQOKLLTVQAQTLASNASPQSPAPSAQISTPCSTQKVKRS 618

QY 424 RLLASQSGAPLLSTPKRBMV-----LMTVEEKLEIERLTKQKELEAKMLAQ 475

DB 619 ILAGN-----CFNTPDSKRAAENAQVRDLQKRVKAMEABIEKMK-KEHLLQK---Q 667

QY 476 KAE 478

DB 668 KDE 670

RESULT 14

US-10-425-115-332919

Sequence 332919, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 332919

LENGTH: 783

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(783)

OTHER INFORMATION: unsure at all xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577\_66737C.1.pap

US-10-425-115-332919

Query Match

Best Local Similarity 23.1%; Score 571; DB 17; Length 783;

Matches 163; Conservative 91; Mismatches 195; Indels 36; Gaps 12;

QY 8 KIGA-TRRPPPARVVRVLRPFVDGTAGASDPPCVGRGMDSCSLEIANWRNHQETLK--- 63

DB 197 EVGAGTAAEAESRILVFLRPMRSRKEAGSRSCVKIVNRKVFLETESSENDVRLRKR 256

QY 64 -----YQDAFYGERSTQDIIYAGSVQPIRLHLEGONASVLAYGPTGAGKTHMLGSP 118

DB 257 GRDSHFCDVSFPDSTTQAEVYSTADLVVEGLQGRNGTVFCYATGAGKTYTMLGTME 316

QY 119 QPGVIPRALMDLLQLTREGAERGPWALSVMYSLEYIOEKVLDLDPASGDLVIREDCR 178

DB 317 NPGVMVLAINDLFSKVTQKH-----SIKLSLEYINETVRDLSLPCR-PLLRDKQ 368

QY 179 GNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLNORSRSHAVLLVKVDQRELAP 238

DB 369 G-IVAAGLTQRSVYSTDEVMEQLQKGNKRTTEPTRVNETSSRSHAVLQVVVEYRSLDGV 427

QY 239 PRQREGKLYLDLAGSDNRRTGKGLRKESGAINTSFLVLGKVDALNOGLPRVPYRD 298

DB 428 NVKRAAGKSLTDLAGSERALATDQRTQSIIEGANINRSLALSSCINALVEGKHIPYRN 487

QY 299 SKLTRLODSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKEV---INRPTNESL- 354

DB 488 SKLTQLLKDSLGGACNTVMIANISPSNLSFGETQNTLHWADRAKEIKTKQOTTVNEEVL 547

QY 355 -QHALGPVKLS-QKE--LLGPPPEAKRARGPEEIEIGSPPEMAAPASAKLSPLQKLS 410

DB 548 DQDSETMLVLELOKENVLRLEQAK-----QQOKLITAEQAQLLAKTSPPSPAPSHV 602

QY 411 MDPAMLERLLSLDRLLASQSGAPLLS-----TPKRERMVLMKTVEKDLIERLTKQK 466

DB 603 STPGSTQKTRRSILAGGGNYFSKLSKRAADNAQVRELQKRVSTLSEIEKMK-KEH 661

QY 467 ELEAK 471

DB 662 LLQIK 666

RESULT 15

US-10-425-114-59708

Sequence 59708, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 59708

LENGTH: 854

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB3689-256-H1\_FLI.pap

US-10-425-114-59708

Query Match

Best Local Similarity 23.1%; Score 571; DB 15; Length 854;

Matches 163; Conservative 91; Mismatches 195; Indels 36; Gaps 12;

QY 8 KIGA-TRRPPPARVVRVLRPFVDGTAGASDPPCVGRGMDSCSLEIANWRNHQETLK--- 63

Db 268 EVGAGTAEAERILVFLRPMSEKEAGSRSCKIVNRKEVFLTESASENDYLRKR 327  
QY 64 -----YQDAFYGERSTQDIYAGSVQPIRLHLLBGONASVLAYGPTGAGKTHMLGSP 118  
Db 328 GRDSHPCFDSVFPDSTTQAEVYSTADLVEGVLOGRNGTFCYGATGAGKTYTMLGTME 387  
QY 119 QGVIPRALMDLLOLTREBAGCRPWALSVTMSYLEIYQEKVLDLDDPASGDLVIREDCR 178  
Db 388 NFGVMVLAINDLFSKVTQKNH-----SILSYLEIYNETVRDLSLSPGS-PLNLRDQK 439  
QY 179 GNILIPGLSQKPTSSPADPERHFLPASRNRITVGATRLNQSRSHAVLIVKVDQERLAP 238  
Db 440 G-IVAAGLTQRSVYSTDEVMELLQKGNKRTTEPTFVNETSSRSHAVLQVVVEYRSLDGV 498  
QY 239 FRQREGKLYLIDLAGEDNRTGNKGLRKESGAINTSFLVIGKVVDALNQGLPRVPYRD 298  
Db 499 NVKRACKLSLIDLAGSERALATDQRTQRTGIEGANINRSLALSSCINALVEGKCHIPYRN 558  
QY 299 SKLTRLLQDSLGSASHILIANIAPERRFYLDTVSALNEAARSKEV---INRPFTNESL- 354  
Db 559 SKLTQLLKDSLGCACNTVMIANISPSNLSFGETQNTLHWADRakeIKTKTOOTVNEEVL 618  
QY 355 -OPHALGPVKLS-QKE--LLGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSS 410  
Db 619 DQPDSETMLVLELQENRVLREQLAK-----QQKLLTAEAOQLTSTSPQSPAPPSHV 673  
QY 411 MDPAMLERLLSLDRLLASOGSOGAPLLS-----TPKRRMVLMTVEEKDLEIERLKTOK 466  
Db 674 STPGSTQRTRRSILAAGGNGVPSKLDKSRHAADNAQVRELQKVKSTLESEIEKMK-KEH 732  
QY 467 ELEAK 471  
Db 733 LLOLK 737

Search completed: November 5, 2004, 19:09:49  
Job time : 83.6338 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 14.1983 Seconds  
(without alignments)  
3300.235 Million cell updates/sec

Title: US-10-797-893-8  
Perfect score: 2472  
Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKOLAQKAEKENHCPTM 487  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2014	81.5	665	2	S62328
2	614	24.8	664	2	T48258
3	612	24.8	784	1	A55236
4	597.5	24.2	703	2	A86319
5	579.5	23.4	1226	2	T15167
6	567	22.9	1254	2	T18277
7	566	22.9	1225	2	A56514
8	565	22.9	1058	2	T47525
9	563.5	22.8	701	1	B44259
10	563	22.8	747	1	A57107
11	561.5	22.7	786	2	A53939
12	561.5	22.7	1056	2	H84777
13	560.5	22.7	1006	2	T02017
14	560.5	22.7	1076	2	B84687
15	560	22.7	885	2	D86151
16	558	22.6	1231	2	A54803
17	557.5	22.6	1388	2	T30335
18	556.5	22.5	1066	1	A48669
19	556	22.5	699	1	S38992
20	551	22.3	2954	2	T14156
21	545	22.0	1229	2	T48959
22	544.5	22.0	935	2	T51930
23	543	22.0	928	2	T10164
24	542	21.9	1070	2	T06733
25	542	21.9	1695	2	A56921
26	541.5	21.9	784	2	T40594
27	541.5	21.9	793	2	JC5831
28	540.5	21.9	929	2	T51932
29	536	21.7	1022	2	E84792

RESULT 1  
S62328  
kinesin-like DNA binding protein KID - human  
C:Species: Homo sapiens (man)  
C:Date: 23-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 16-Aug-2004  
C:Accession: S62328  
R:Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Tsukita, S.; Inoue, .  
EMBO J. 15, 457-467, 1996  
A:Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes an  
A:Reference number: S62328; MUID:96174806; PMID:8599929  
A:Accession: S62328  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-665 <TK>  
C:Superfamily: kinesin motor domain homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:44-374/Domain: kinesin motor domain homology <KNOT>  
F:128-135/Region: nucleotide-binding motif A (P-loop)  
Query Match 81.5%; Score 2014; DB 2; Length 665;  
Best Local Similarity 82.4%; Pred. No. 1e-128;  
Matches 416; Conservative 3; Mismatches 44; Indels 42; Gaps 4;

## ALIGNMENTS

30	533	21.6	742	1	S58691	kinesin-related pr
31	532	21.5	1056	1	G02157	kinesin-like spind
32	529.5	21.4	843	2	S44868	kinesin heavy chai
33	525.5	21.3	1060	1	A40264	kinesin-related pr
34	525	21.2	1150	1	A52889	kinesin-like prote
35	522.5	21.1	805	2	S64238	kinesin-related pr
36	520.5	21.1	1031	1	A38713	kinesin heavy chai
37	520.5	21.1	1067	2	S33417	kinesin-like prote
38	519	21.0	967	1	A35075	kinesin heavy chai
39	518.5	21.0	958	2	T20621	hypothetical prote
40	518	21.0	1584	1	JN0114	kinesin-related pr
41	518	21.0	1584	2	T15822	kinesin-like prote
42	517	20.9	2663	1	S28261	centromere protei
43	516	20.9	813	2	T46242	kinesin-like prote
44	512	20.7	1027	2	S37711	kinesin heavy chai
45	511	20.7	1121	2	T06065	hypothetical prote

Qy 2 GRCLSKIGATRRPPPARVAVRLRPVDCGTAGASDPPCVRGWMDSCSLEITANWRNHQET 61  
Db 26 GRCLSKIGATRRPPPARVAVRLRPVDCGTAGASDPPCVRGWMDSCSLEITANWRNHQET 85  
Qy 62 LKQFDAPYGERSTQDDIYAGSVQPIRLHLLLEGQNAS-VLAYGPTGAGKTHMLGSPROP 120  
Db 86 LKQFDAPYGERSTQDDIYAGSVQPIRLHLLLEGQNAKVVLAYGPTGACKT-THAQOPRAT 144  
Qy 121 GVTPRALMDLLQLTRREGAERPWALSVTMSYLIYQKVLDDLLDPAGSDLVIREDCRGN 180  
Db 145 WGDPAAGHGPPAAHKGGCGRPGAMGLSVTMSYLIYQKVLDDLLDPAGSDLVIREDCRGN 204  
Qy 181 ILIPLGSKPTSSPADFERHFLPASRRTVCATRLNQRSSRSHAVLLVKVQDRERLAPFR 240  
Db 205 ILIPLGSKPTSSPADFERHFLPASRRTVCATRLNQRSSRSHAVLLVKVQDRERLAPFR 264  
Qy 241 QREGLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVKGVVDALNQGFLPRVYRDSK 300  
Db 265 QREGLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVKGVVDALNQGFLPRVYRDSK 324  
Qy 301 LTRLLQSLGSSAHSILIANIAPRRFYLDVTVALNFAARKEVINRPFTHESLOPHALG 360  
Db 325 LTRLLQSLGSSAHSILIANIAPRRFYLDVTVALNFAARKEVINRPFTHESLOPHALG 384  
Qy 361 PVKLSQKELGPPPAKARGPEEEIGSPPEMAA-----PASASQ 400  
Db 385 PVKLSQKELGPPPAKARGPEEEIGSPPEMASSLCCLPQTPPTPEAKAAWTRPCGAPP 444  
Qy 401 KLSPLQKLSSMDPAMLERLLSLDLLASQSQGAPLLSTPKRERMVLMKTVEEKDLEIER 460

Db 445 QLGPSACL-----PGEPCALLSTPKRVMVLMKTVEEKDLEIR 484

QY 461 LXTKQKELEAKMLAQAEKENHCP 485

Db 485 LXTKQKELEAKMLAQAEKRTIVP 509

RESULT 2

T48258

kinesin-like protein - Arabidopsis thaliana

N;Alternate names: protein T1E22.130

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence [revision 20-Apr-2000 #text\_change 09-Jul-2004

C;Accession: T48258

R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, March 2000

A;Reference number: 224489

A;Accession: T48258

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-664 <BEV>

A;Cross-references: UNIPROT:Q9L288; EMBL:AL162874

A;Map position: 5

A;Experimental source: cultivar Columbia; BAC clone T1E22

C;Genetics:

A;Description: may be part of a motor protein that provides anterograde fast axonal tra

A;Superfamily: kinesin-related protein KIF3; kinesin motor domain homology

C;Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop

C;F10-349/Domain: head globular #status predicted <HGL>

F120-350/Domain: kinesin motor domain homology <KMT>

F106-113/Region: nucleotide-binding motif A (P-loop)

F350-580/Domain: helical rod #status predicted <ROD>

F581-784/Domain: tail globular #status predicted <TGL>

F112/Binding site: ATP (lys) #status predicted

Query Match 24.8%; Score 614; DB 2; Length 664;

Best Local Similarity 32.1%; Pred. No. 8.2e-34; Indels 64; Gaps 13;

Matches 163; Conservative 94; Mismatches 187;

QY 18 ARVRVAVLRPFV-----DGTAGADPPCVRGMDSCSLEIANRNRHQTLEK---YQFDA 68

Db 78 SNRVVLVRVLPFLPREISDSCDGRSCVSVIGDGGDTSEVAVYKDPDSCRNESYQLDA 137

QY 69 FYG-ERSTQODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHMLGSPQGVIPRAL 127

Db 138 FYGREDNNVKHIFDREVSPLPIGFHGFNATVLAYGATGSGKFTTQGIDELPLMLPTM 197

QY 128 MDLLQLTREGAEGRPWALSVTMSYLEIQYOKVLDLDPASGDLVIREDCRGNILIPGLS 187

Db 198 STILSMCEKTRSAE-----TSYEVYMDRCWDLLEVKDNEIAVWDKDGQVHLKGLS 250

QY 188 QKPTISSPADFERHFLPASRNTVGATRLNQRSSSHAVLVKVDQRLAPFRQREGKLY 247

Db 251 SVPVKSMSFEQAYLCGVQRRKVAHTGLNDVSSRSHGVLVISVTSQGLVT-----GKIN 304

QY 248 LIDLAGEDNRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPPYRDSKLTLLQD 307

Db 305 LIDLAGEDNRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVPPYRDSKLTLLQD 364

QY 308 SLGGSASHILTIANTAPERFRFYDTVSALNFAARKEVINRPFNTNESLQPHALGP-VKLS- 365

Db 365 SLGGSASHILTIANTAPERFRFYDTVSALNFAARKEVINRPFNTNESLQPHALGP-VKLS- 418

QY 366 QKELLGPPEAKRGPBEEETGSPPEMAAPASAKLSPLQKLSMDPAMLE----- 417

Db 419 EAKLQAWLESKGMKSAHRMAIRSPMLGTNTQTSISQSSVKKLCHRSAIAESAKLAGTG 478

QY 418 -----RLSLDLRLIASQSQGAPL-----LSTPKRE-----RMLVLMKTVEEK 454

Db 479 QRDFAVTARNLFGVETTLAASH--LWEPINQLASPTKEDRDTSGREENLLVSEASLRDN 536

QY 455 DLEIRLTKQKELEAKMLAQAEKEN 482

Db 537 TLDVEKKYTELSPLREALSPIDSNAPKN 564

RESULT 3

A55236

kinesin-related protein KLP68D - fruit fly (Drosophila melanogaster)

N;Alternate names: kinesin-like protein 5; KLP5

C;Species: Drosophila melanogaster

C;Date: 23-Mar-1995 #sequence [revision 23-Mar-1995 #text\_change 09-Jul-2004

C;Accession: A55236; E41298

R;Pesavento, P.A.; Stewart, R.J.; Goldstein, L.S.B.

J. Cell Biol. 127, 1041-1048, 1994

A;Title: Characterization of the KLP68D kinesin-like protein in Drosophila: possible ro

A;Reference number: A55236; MUID:95050960; PMID:7525600

A;Accession: A55236

A;Molecule type: mRNA

A;Residues: 1-784 <PES>

A;Cross-references: UNIPROT:P46867; GB:U15974; NID:9595912; PIDN:AAA69929.1; PID:g56509

R;Stewart, R.J.; Pesavento, P.A.; Woerpel, D.N.; Goldstein, L.S.B.

Proc. Natl. Acad. Sci. U.S.A. 88, 8470-8474, 1991

A;Title: Identification and partial characterization of six members of the kinesin supe

A;Reference number: A41298; MUID:92020874; PMID:1924306

A;Accession: E41298

A;Molecule type: DNA

A;Residues: 'TC', 222-337, 'VRGQV' <STE>

A;Cross-references: GB:M74431; NID:g157791; PIDN:AAA28658.1; PID:g157792

C;Genetics:

A;Gene: FlyBase:Klp68D; KLP5

A;Cross-references: FlyBase:FBgn0004381

C;Function:

A;Description: may be part of a motor protein that provides anterograde fast axonal tra

C;Superfamily: kinesin-related protein KIF3; kinesin motor domain homology

C;Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop

C;F10-349/Domain: head globular #status predicted <HGL>

F120-350/Domain: kinesin motor domain homology <KMT>

F106-113/Region: nucleotide-binding motif A (P-loop)

F350-580/Domain: helical rod #status predicted <ROD>

F581-784/Domain: tail globular #status predicted <TGL>

F112/Binding site: ATP (lys) #status predicted

Query Match 24.8%; Score 612; DB 1; Length 784;

Best Local Similarity 33.0%; Pred. No. 1.4e-33; Indels 88; Gaps 18;

Matches 175; Conservative 84; Mismatches 184;

QY 12 TRRP-----PPARVRVAVLRPFVDGTAGADPPCV-----RGMDSCLSIANN--RN 57

Db 5 SRPPTGSSQTNPCVQVVRCPMSNRBSERSPEVNVVYVNRGV---VELQNVDGN 60

QY 58 HOETLYKQPDAPYGRSTOODIYAGSVQPIRLHLEGQNASVLAYGPTGAGKTHM---L 114

Db 61 KEORVFTYDAAYDASATQTTLYHEVFPVLSVLEGGNGCIFAYGQTGTGKTFTMEGVR 120

QY 115 GSPEQGVIPRALMDL-LQLTREGAEGRPWALSVTMSYLEIQYOKVLDLDPASGDLVI 173

Db 121 GNDELMIIPRTPEQIWLHINTEN-----FQFLVDVSVLEIYMEELRDLKPKNSKHLEV 175

QY 174 REDCRGN-ILIPGLSQKPISSPADFERHFLPASRNTVGATRLNQRSSSHAVLVKVDQ 232

Db 176 RE--RSGVYVFNHAINCKSVEDMIKVMQVGNKRTVGTFTNNHSSRSHAFMIKIE 233

QY 233 RELAPFRQREGKLYLIDLAGEDNRTGNKGLRKESGAINTSFLVLGKVVDALNOGLP 292

Db 234 CD-TETNTIKVGNLIDLAGESRQSKTGASERLKEASKINLALSLSLGNVISALAE 292

QY 293 RVPYRDSKLTLLQDSLGGSAHSILTIANTAPERFRFYDTVSALNFAARKEVINRPFNTNE 352

Db 293 HVPYRDSKLTLLQDSLGGSKTIMIANTGPSNVTNETLTTLRYGSRAKSIQNQPIKNE 352

QY 353 SLOPHALGPVKLSQ-----KELLGPPEAKRA-----RGPEEEI----- 386

Db 353 DPQ-----DAKLEYQEEIERLKLIGPQQQQRSEKQVTAQRKVRKPKKETTKEMSDS 407

QY 387 -----GSPEPMAAPASAKLSPLQKLSMDPAMLERLILDLRLIASQSQG 433

Db 408 LQVSTIEQPVEDSDPE---GAESSEKENEAEVAKSNEELERERVEN-SKLAAKLAELE 463

QY 434 APILSTPKRERMVLMKTVEEKLEIER-----LTKQKELEAKMLAQAE 479

Db 464 GQLVRGGN---LLDTYSERQIELEKLVIEAKRKRETEIQOQLQOE 510

## RESULT 4

A86319  
F25116.11 protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: A86319  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A86319  
A:Status: preliminary  
A:Molecule type: DNA  
A:Gene: klp1  
A:Residues: 1-703 <STO>  
A:Cross-references: UNIPROT:Q9FZ77; GB:AE005172; NID:g9795601; PIDN:AAF98419.1; GSPDB:GN  
C:Genetics:  
A:Map position: 1

Query Match	24.2%	Score	597.5	DB 2	Length	703			
Best Local Similarity	33.9%	Pred. No.	1.2e-32						
Matches	173	Conservative	88	Mismatches	172	Indels	77	Gaps	17

Qy 18 ARVAVRLRPFDVGTAGADPPCVRGMDSCSLBIANRNHOETLK-----YQDAF 69  
Db 150 SRLVFLVRPMGKERENGSRCCVKLNKRDVLTFTENDYLRKLVRHFTDSS 209  
Qy 70 YGERSTODIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPQPGVTPRALMD 129  
Db 210 FPETTTQGVYSTTTGDLVEAVLEGRNGSVFCYATGAGKTYTTLGTWENPGVMVLAID 269  
Qy 130 LLQLTREGAEGRPWALSVTMSYLEIYOEKVLDDLPASGDLVIREDCRGNILIPGLSOK 189  
Db 270 LFAKVRORSLDGNH---VHLSYLEVNETVRDLSLSPR-PLIUREDKQ---VMALLQR 321  
Qy 190 PISGFADPFRHFLPASNRRTVGATRLNQRSSRSRAVLVKKVDQRLAPFR--OREGKLY 247  
Db 322 -----GNQRTTEPRCNETSSRSRAILQVIVEYKTRDASNNIISRVCKLS 367  
Qy 248 LIDLAGSDNRRTGKGLRKESGAINTSLFVLGKVDALNQGPRVPYRDSKLTLLQD 307  
Db 368 LIDLAGSERALATDQRTLSLEGANINRSLALSSCINALVEGKHIPYRNSKLTQLLKD 427  
Qy 308 SLGSGSAHSILIANAPERRFYLDTVSALNFAARKEV-INRPFTNESILOPHALGPVKLSQ 366  
Db 428 SLGSGCNTVMIANISPSQSGFQNTLHWADRAKEIRVKEVEVNEEV-----VQVGE 480  
Qy 367 KELIGPPEAKRARGPEEEIGSPFMAAPASASOKLSPLQKLS-----SMDPAML 416  
Db 481 EE--GAQAKLLLELQXE---NSELRVQLAKQQQKLTQAEINAAANNNNISLTTPSI 535  
Qy 417 ERLLSLDRLLASQSGQA--PLLS-----TP-----KRMV--LMKTVEKDLIERLKL 462  
Db 536 SSLMTPTPSALTAQOKKPRHSLSGCTPTESLARTKAEAEVAKELQTLVKALMEMERMK 595  
Qy 463 -----TKOK-ELEAKWLAKAEKENHC 484  
Db 596 REHGLQMKQKQDLMDLCLSRXSEKTPERC 625

## RESULT 5

151617  
kinesin-like protein 1 - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Aug-2004

C:Accession: I51617; A48835; S48837  
R:Vernos, I.; Raats, J.; Hirano, T.; Heasman, J.; Karsenti, E.; Wylie, C.  
Cell 81, 117-127, 1995  
A:Title: Xklp1, a chromosomal Xenopus kinesin-like protein essential for spindle organ-  
A:Reference number: A56221; MUID:95236444; PMID:7720067  
A:Accession: I51617  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1226 <VER>  
A:Cross-references: UNIPROT:Q91784; EMBL:X82012; NID:g562792; PIDN:CAA57539.1; PID:g56-  
R:Vernos, I.; Heasman, J.; Wylie, C.  
Dev. Biol. 157, 232-239, 1993  
A:Title: Multiple kinesin-like transcripts in Xenopus oocytes.  
A:Reference number: A48835; MUID:93246065; PMID:8482413  
A:Accession: A48835  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 9-162, 'L', 164-338 <VE2>  
A:Experimental source: oocyte  
A:Note: sequence extracted from NCBI backbone (NCBIP:130975)  
C:Genetics:  
A:Gene: klp1  
C:Superfamily: kinesin motor domain homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:9-343/Domain: kinesin motor domain homology <RMOT>  
F:87-94/Region: nucleotide-binding motif A (P-loop)

Query Match	23.4%	Score	579.5	DB 2	Length	1226			
Best Local Similarity	32.1%	Pred. No.	4.3e-31						
Matches	157	Conservative	95	Mismatches	172	Indels	65	Gaps	13

Qy 20 VRVAVRLRPFDVGTAGADPPCVRGMDSCSLBIANRNHOETLK-----KYQDFAPYGRS 74  
Db 9 VRVALRCPLVP-----KENEGCKMCLTFVQEQOVIIVGTEKSFYDYVDFDSA 58  
Qy 75 TQDIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGS-----PEQP--GVIPRAL 127  
Db 59 EQEVNYSVAAPLIKGLFKGNATVAYGQTGSGKTYSMGAYTHNQENETVGVIPRTV 118  
Qy 128 MDLLQLTREGAEGRP-WALSVTMSYLEIYOEKVLDDLPDPA---SGDLVIREDCRGNIL 183  
Db 119 IALFREIHQ-----RPEWENFLKVSYLEIYNEEILDLLYAADKNTTISIREPKGKIKI 173  
Qy 184 PGLSQKPISSFADPFRHFLPASNRRTVGATRLNQRSSRSRAVLVKKVDQRLAPFRQRE 243  
Db 174 CGLTERDVKTALDTLSCLEQNSRSTVASTAMNSQSSRSRAIFTISIEQRKEGDKNSFR 233  
Qy 244 GKLYLIDLAGSDNRRTGKGLRKESGAINTSLFVLGKVDALNQGPR---VPYRDSK 300  
Db 234 SKLHLVDLAGSERQKTKAEGDLKEGISINRGLLCLGNVISALGDESKGGFVYRDSK 293  
Qy 301 LTRLLQDSLGSAHSILIANAPERRFYLDTVSALNFAARKEVINRPFTHESLOPHALG 360  
Db 294 LTRLLQDSLGSGNSHTLMIACVSPADSNMEETLNTLYADRAKIKNPKIVNTDFQAAELQ 353  
Qy 361 PVKLSOKELLGPPEAKRARGPEEEIGSPFMAAPASASOKLSPLQKLSMDPAMLERLL 420  
Db 354 RLKLQVQEL--QVLLLOAHGGLTPVNSMEPSNLQSLMERNKVKLENG-----401  
Qy 421 SLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLIERLKTOKELE-----AKM 472  
Db 402 KLSRELGEAAVQTAQFL-----EKIIMTEQONEK-----LGSKWEELKQHAACKVNQIR 450  
Qy 473 LAQKAEKE 481  
Db 451 LVETLEDOE 459

## RESULT 6

151617  
kinesin heavy chain - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

Fri Nov 12 12:26:29 2004

C:Accession: T18277  
 R:deHostos, B.L.; McCaffrey, G.; Vale, R.D.  
 Submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z18853  
 A:Accession: T18277  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1254 <DEH>  
 A:Cross-references: UNIPROT:Q94463; EMBL:U41289; MID:g1526990; PID:g1526991; PIDN:AA8077

Query Match 22.9%; Score 567; DB 2; Length 1254;  
 Best Local Similarity 36.1%; Pred. No. 3.1e-30;  
 Matches 126; Conservative 79; Mismatches 132; Indels 12; Gaps 5;

QY 15 PPARVAVRVRPPFVDTAGASDPPVGRGMDSCSLEI-ANRNHQTETLYQYDAFYGER 73  
 DB PVSINRVCVRPLTELEKGRNEHSIVHFFDSKISIRANGP-----QFTDFRIFGVQ 77

QY 74 STQODIYAGSVQPIRLHLEGQNASVLAGPTGAGKTHMTLMSGPPQGVIPRALMDL--- 130  
 DB ETQSQIFEDVAEPIVNDLFDGVHTIAYGQTASGKFTTWGDPDPSHGIIIPRVIESIFVG 137

QY 131 LQUTREGAEGRPWALSVTMSYLEIYQEKVLDLDDPASGLVIREDCRGNILIPGLSOKP 190  
 DB ISKREKDT-SLSLAFCLKISALELYNEKLYDYIASKNLNIREHKQGIYVEGISEIV 196

QY 191 ISSFADFERHFLPASRNRTVGATRLNORSRSHAVLLVQDQORERLAPRQREGKLYLID 250  
 DB ITSIEAYNPLNINNNRAISTKMSAASRSHSVLMISLQ-QNLSMESSKISKLFLVD 255

QY 251 LAGSEDRRTGKGLKSGAINTSLFVGVKVVDMALNOGLPRVPRYRDSKLTRELQDSIG 310  
 DB LAGSDISHTGAEGRMQEAKNINLSLALGKVINALTGCVNYPVYRDSKLTRELQDSIG 315

QY 311 GSAHSILIANIAPERRFLDTVSAALNFAARKEVINRPTNESLQPHAL 359  
 DB GNSKTSLIINCSNNNEHETITLQGTAKTIKQPKINKKITYHEL 364

RESULT 7  
 A56514  
 Chromokinesin - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 09-Jul-2004  
 C:Accession: A56514; I50691  
 R:Wang, S.Z.; Adler, R.  
 J. Cell Biol. 128, 761-768, 1995  
 A:Title: Chromokinesin: a DNA-binding, kinesin-like nuclear protein.  
 A:Reference number: A56514; MUID:95181533; PMID:7876303  
 A:Accession: A56514  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1225 <WAN>  
 A:Cross-references: UNIPROT:Q90640; GB:U18309; MID:g603760; PIDN:AAC59666.1; PID:g603761  
 R:Wang, S.Z.; Adler, R. U.S.A. 91, 1351-1355, 1994  
 Proc. Natl. Acad. Sci. U.S.A. 91, 1351-1355, 1994  
 A:Title: A developmentally regulated basic-leucine zipper-like gene and its expression  
 A:Reference number: A53451; MUID:94151328; PMID:8108415  
 A:Accession: I50691  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 728-1086, 'RI' <WA2>  
 A:Cross-references: EMBL:U04821; MID:g440792; PIDN:AAAL8960.1; PID:g440793  
 C:Genetics:  
 A:Gene: sw3-3  
 C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology  
 C:Keywords: ATP; DNA binding; mitosis; nucleotide binding; P-loop  
 F:11-344/Domain: kinesin motor domain homology <KMT>  
 F:88-95/Region: nucleotide-binding motif A (P-loop)

Query Match 22.9%; Score 566; DB 2; Length 1225;  
 Best Local Similarity 31.3%; Pred. No. 3.5e-30;  
 Matches 159; Conservative 103; Mismatches 182; Indels 64; Gaps 14;

QY 17 PARVAVRVRPPFV-DGTAGA-----SDPPVGRGMDSCSLEIANRNHQTETLYQY 66  
 DB PVRV---VRCRPLVPKETSEGCQCLSLFVPGEPQVIVGSDKA-----FTY 51

QY 67 DAFYGERSTQODIYAGSVQPIRLHLEGQNASVLAGPTGAGKTHMTLMSGPPQGVIPRALMDL--- 119  
 DB DYVDPSVEQEVVENTAVPIRIGFKGNATVLAGYGTGSGKTYSGMGTYTASQEHDP 111

QY 120 PGVIPRALMDLLOLTREGEAEGRPWALSVTMSYLEIYQEKVLDLDDPA---SGDLVIRE 176  
 DB MGVIPIR---VIKLFKEKEQDWEFVLKVSYLEIYNEDILDLCCSRERSQISIRE 167

QY 177 CRGNILIPGLSOKPISFPADFERHFLPASRNRTVGATRLNORSRSHAVLLVQDQORERL 236  
 DB PKEGKIVGLTERNVASARDTVSCLEQGNCRVTASTAMNSQSSRSHAFITCIDQKK 227

QY 237 APPRQREGKLYLIDLAGSEDRRTGKGLKSGAINTSLFVGVKVVDMALNOGLPR--- 293  
 DB DKNSSFHKLHLVDLAGSERQKTKAEGDLKEGININRGLLCLGNVISALGEENKKG 287

QY 294 VPRYRDSKLTRELQDSIGSAHSILIANIAPERRFLDTVSAALNFAARKEVINRPTNES 353  
 DB VPRYRDSKLTRELQDSIGSHTLMIAVCSPADSNLEETLNTLYADRAKIKNKDIV 347

QY 354 LQPHALGPVK--LSQKELL-----GPPEAKRARGPEEEIGSPPEMAAPASASQKLS- 403  
 DB POAAELNHLKQVQVQLVLLQAHGGTLPVSNWAPSENLOSLMEKQSLMEENEKLS 407

QY 404 PQKLSMDPAMLERLLSLDRLLASQSGQAPLL---STPKRERMVLMKTVEKDLLE--IE 459  
 DB GLSEAAQTQAMLEIRIIVTEQENRKNAKLEQLQQAHAVCKLDLQKLETVDEELKEN 467

QY 460 RLTKQKEL-----EAKMLAQKAEKEN 482  
 DB VIRNLQVLAQFQSESAAAAAATAEMAN 495

RESULT 8  
 T47525  
 kinesin-related protein-like - Arabidopsis thaliana  
 N:Alternate names: protein F16L2.60  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
 C:Accession: T47525  
 R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.;  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24468  
 A:Accession: T47525  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1058 <JOR>  
 A:Cross-references: UNIPROT:Q9LZU5; EMBL:AL162459  
 A:Experimental source: cultivar Columbia; BAC clone F16L2  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 56/2; 99/3; 151/3; 310/3; 343/2; 395/3; 419/1; 444/3; 462/3; 489/3; 527/1; 5  
 A:Note: F16L2.60  
 C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology.

Query Match 22.9%; Score 565; DB 2; Length 1058;  
 Best Local Similarity 29.5%; Pred. No. 3.4e-30;  
 Matches 156; Conservative 99; Mismatches 201; Indels 72; Gaps 12;

QY 5 RLSKIGATRRPPPAR---VRVAVRLRPPFVDGTAGASDPPVGRGMDSCSLEIANRNHQTET 61  
 DB RSSESNTNRNKEKGVNQVILRCPLSEDEARIHTPVVI-----SCN-----ENREV 80

QY 62 LKQY-----FDAPYGERSTQODIYAGSVQPIRLHLEGQNASVLAGPTGAGK 109  
 DB AATQSTAGKHIDRFADFVKVFPASQKLDYDQAIPIVFEVLEGNCTIFAYGQTGTGK 140

QY 110 THTMLGS-----PEQPGVIPRALMDLLOLTREGEAEGRPWALSVTMSYLEIYQEKV 161



Db 141 TYTMEGGARKKNGEFPDAGVPIRAVKQIFDLEAQGAEE-----YSMKVTFLELYNEEIS 195  
Qy 162 DLLDP-----ASGDLVIREDCRGNILIPGLSOKPISSFADFERRHFLPASRRNRTVG 211  
Db 196 DLLAPEETIKFVDEKSKSIAMMEDGKSVFVRGLSEIEIVSTANEIYKILEGSAKRTA 255  
Qy 212 ATRLNQSSRSRSHAVLLVKVDQERLAPRQ--REGKLYLIDLAGSEDRNRTGNKGLRLKE 269  
Db 256 ETLNKKOSSRSRSHSIFSIHIKENTPEGEEMIKCKGLNVLDSGSENISRSRGAREGRARE 315  
Qy 270 SGANTSIFVLGKVVADALNQLPVPYRDSKLTLLQDSLGSAHSLIANIAPERFYL 329  
Db 316 AGEINKSLTUGRVINALVHSGHIPYRDSKLTLLRESLGGTKTCTVIATISPSIHCL 375  
Qy 330 DTVSALNFAARSKVINRPFNTESLQHALGVKLSOKELGPPPEAKRARGPEEEETGSP 389  
Db 376 ETLSTLOYAHRANKIKKPEINQMKSVAMKDIYSIDRL--KQEVVAAR--EKNGIYIP 432  
Qy 390 EPMAAPASAKQ--LSPLQKLSMDPAMLERLLSLDRLLASQSGGAPLLSTPKRERMVL 447  
Db 433 KDRYIQEAEKKAEMAKIERLELQSESKRQVVDLQELVNSQQILTAEELSEKLEKTEK 492  
Qy 448 MK-----TVSEKDLIERIKTKQKELEAKMLAQKAE 478  
Db 493 ETEHSLFDLEEKYRQANATIKKEFVISNLLKSEKSLVERAFQLRTE 540

## RESULT 9

B44259  
Kinesin-related protein KIF3A - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 09-Jul-2004  
C:Accession: B44259; S27872  
R:Alzawa, H.; Sekine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.  
J. Cell Biol. 119, 1287-1296, 1992  
A:Title: Kinesin family in murine central nervous system.  
A:Reference number: A44259; MUID:93077686; PMID:1447303  
A:Accession: B44259  
A:Molecule type: mRNA  
A:Residues: 1-701 <ATZ>  
A:Cross-references: UNIPROT:P28741; EMBL:D12645; NID:G220469; PIDN:BAA02166.1; PID:G2204  
A:Experimental source: brain  
A:Note: sequence extracted from NCBI backbone (NCBIP:118911)  
C:Complex: heterodimer with KIF3B (PIR:A57107); the KIF3A/3B heterodimer associates with  
C:Function:  
A:Description: KIF3 complex is a motor protein that provides anterograde fast axonal tra  
C:Superfamily: kinesin-related protein Kif3; kinesin motor domain homology  
C:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding; nucleotide  
F:1-368/Domain: head globular #status predicted <HGL>  
F:15-351/Domain: kinesin motor domain homology <KMOT>  
F:100-107/Region: nucleotide-binding motif A (P-loop)  
F:369-599/Domain: helical rod #status predicted <ROD>  
F:600-701/Domain: tail globular #status predicted <TGL>  
F:106/Binding site: ATP (Lys) #status predicted

Query Match 22.8%; Score 563.5; DB 1; Length 701;  
Best Local Similarity 30.2%; Pred. No. 2.3e-30;  
Matches 169; Conservative 83; Mismatches 184; Indels 123; Gaps 16;

Qy 9 IGATRRPPPA-RVRVAVRLRPF-----VDGTAGASDPPPCVRG 44  
Db 3 INKSEKPESCDNKVVVRCRPLNREKSMCVQAVSDVEMRGITVHKTDSSNEPP----- 58  
Qy 45 MDSCSLIANWRNHQETLKQYQDAFYGERSTQDDIYAGSVQPIRLHLLLEGONASVLA YGP 104  
Db 59 -----KTFTFDTVFGPESKQLDVNLTARPIIDSVLEGYNGTIFAYGQ 101  
Qy 105 TGAGKTHMLGSPPEQ---GVIPRALMDLL-QLTRREGAERGPWALSVTMSYLEIYOEKV 160  
Db 102 TGGKPTMEGVRAVPLGRLGVIPNSFAHIFGHAKAEG-----DTRFLVRVSYLEYINEEV 157  
Qy 161 LDLL-DPASGDLVIREDCRGNILIPGLSOKPISSFADFERRHFLPASRRNRTVGATRLNQRS 219

Db 158 RDLGKQDTQRLVKERPDVGVIKDLISAYVVVNADMDRITMTLGHKNRSVGATWNNEHS 217  
Qy 220 SRSHAVLLVKVDQERLAP--FRQEGKLYLIDLAGSEDRNRTGNKGLRLKESGAINTS 277  
Db 218 SRSHAITTTITTECEKGVGMHVMGKLHLVLDLAGSERQAKTGATGQRLKEATKINLS 277  
Qy 278 FVLGKVVADALNQG-LPRVPYRDSKLTLLQDSLGSAHSLIANIAPERFRLYDVTVSALN 336  
Db 278 STLGNVISALVDGKSTHVPYRNSKLTLLQDSLGSGNKTWMCANIGPADYNYDEITSLR 337  
Qy 337 PAARSKVINRPFNTESLQHALGV-----KLSQ-KELGLGPPPEAKRARGPEEBEI 386  
Db 338 YANRAKNIKNAKINEDPKOALLRQFQKEIBELKKLEEGEVSGSDISGSEEDDEGEL 397  
Qy 387 G-----SPEPMA-----APASASQKL 402  
Db 398 GEDGEKKRRDQAGKKVSPDKMVENQAKTIDERKALETKLDMEEEERNKARAELERR 457  
Qy 403 SPLQKLSMDPAMLERLLSLDRLLASQSGGAPLLSTPKRERMVLMTVEEKDLIERIK 462  
Db 458 KDLKAAQOEHSLEKLSALEKKVI---VGVVDLLAKAEEQE---KLLEESNWELE--- 507  
Qy 463 TKQKELEAKMLAQKAEKE 481  
Db 508 --ERRRRAEQLRKEEKE 524

## RESULT 10

A57107  
Kinesin-related protein KIF3B - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: A57107  
R:Yamazaki, H.; Nakata, T.; Okada, Y.; Hirokawa, N.  
J. Cell Biol. 130, 1387-1399, 1995  
A:Title: KIF3A/B: a heterodimeric kinesin superfamily protein that works as a microtub  
A:Reference number: A57107; MUID:96032268; PMID:7559760  
A:Accession: A57107  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-747 <YAM>  
A:Cross-references: UNIPROT:Q61771; GB:D26077; NID:g1060922; PIDN:BAA05070.1; PID:g1060  
A:Experimental source: brain  
C:Complex: heterodimer with KIF3A (PIR:B44259); the KIF3A/3B heterodimer associates wi  
C:Function:  
A:Description: KIF3 complex is a motor protein that provides anterograde fast axonal tr  
C:Superfamily: kinesin-related protein Kif3; kinesin motor domain homology  
C:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding; nucleoti  
F:1-363/Domain: head globular #status predicted <HGL>  
F:10-346/Domain: kinesin motor domain homology <KMOT>  
F:98-103/Region: nucleotide-binding motif A (P-loop)  
F:364-592/Domain: helical rod #status predicted <ROD>  
F:594-747/Domain: tail globular #status predicted <TGL>  
F:102/Binding site: ATP (Lys) #status predicted

Query Match 22.8%; Score 563; DB 1; Length 747;  
Best Local Similarity 30.7%; Pred. No. 2.8e-30;  
Matches 169; Conservative 95; Mismatches 189; Indels 98; Gaps 16;

Qy 20 VRVAVRLRPFVDGTAGASDPPPCVRGMD---SCSLEIANWRNHQETLKQYQDAFYGERST 75  
Db 10 VVVVRCRP-MNGKKAASVDKVDVVKLGQSVKPKGTSHMPKTFDFAVDYDNNAK 68  
Qy 76 QDDIYAGSVQPIRLHLLLEGONASVLA YGPTGAGKTHTM---LGSPEQGVIPRALMDLL- 131  
Db 69 QFELYDETFRVDSVLQGFNGTIFAYGQTGTGTGTMEGVRGDPKRGVIPNSFDHIFT 128  
Qy 132 QLTRREGAERGPWALSVTMSYLEIYOEKVLDD-LPASGDLVIREDCRGNILIPGLSOKP 190  
Db 129 HISRSQNNQ-----YLVRSYLEYIQEIRDLISKOQTKRELKELKRPDPTGVYVVDLSFV 183  
Qy 191 ISSFADFERRHFLPASRRNRTVGATRLNQRSRSHAVLLVKVDQERLAPFRQ--REGKLYL 248

```

Db      184  TKSVEIEHVMVGNQNRSGATNMNEHSRSHAFVITIECEVGLDGENHVRGKJNL 243
QY      249  IDLAGSEDRNRRTGNKGLRLKESGAINTSFLVLGVVDALNQG-LRPVPRYDSKLTLLQD 307
Db      244  VDLAGSERQAKTGAQGERLKEATKINLSALGNVISALVDGKSTHIPPYRDSKLTLLQD 303
QY      308  SIGGSAHSILIANIAPERRPYLDTVSALNPAARSKVINRPFNE----- 352
Db      304  SLGGNAKTVMANVNGPASYNVEETLTTRYANRAKNIKPKRVNEDPKDALLREFOEETA 363
QY      353  ----SLQPHALGPVKLSQKELL-----GPEAKRARGPEEEETG----- 387
Db      364  RLKAQLEKSGIRKREKRREGGGGGGEBEGBEGBEEDGDDKDDYWRBQKLEI 423
QY      388  ----SPEPMAAPASQKSLPLQK-----LSSMDPAMLERLLSLDLLAS 428
Db      424  EKRAIVEDHSLVAERKRLKEKEKWEDLRREKDAEMLGAKIKAMESKLLVGGKNIVD 483
QY      429  QSQGAPLL-----STPKRRMV--LMKTVEKDLER-----LKTQK-KE 467
Db      484  HTNEQQKILEQKQKEIAEQKRREIRIQQOMESRDEETLELKETYTSLOQOEVDIKTKLKK 543
QY      468  LEAKMLAQKAE 478
Db      544  LFSKLQAVKAE 554

RESULT 11
A53939
kinesin homolog KHP1 - Chlamydomonas reinhardtii
C;Species: Chlamydomonas reinhardtii
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: A53939
R;Walther, Z.; Vashishtha, M.; Hall, J.L.
J. Cell Biol. 126, 175-188, 1994
A;Title: The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous protein.
A;Reference number: A53939; MUID:94299638; PMID:8027176
A;Accession: A53939
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-786 <WAL>
A;Cross-references: UNIPROT:P46869; EMBL:L33697; NID:g497696; PID:g4976
A;Note: authors translated the codon AAC for residue 753 as Asp
C;Genetics:
A;Gene: FLA10
C;Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C;Keywords: ATP; coiled coil; nucleotide binding; P-loop
F;11-359/Domain: kinesin motor domain homology <KNOT>
F;97-104/Region: nucleotide-binding motif A (P-loop)

Query Match      22.7%; Score 561.5; DB 2; Length 786;
Best Local Similarity 30.7%; Pred. No. 3.8e-30;
Matches 172; Conservative 88; Mismatches 190; Indels 111; Gaps 19;

QY      16  PPA-----RVRAVRLRPP-----VDGTAGASPPPCVRGMDSCSLSIANWR-NHQITLK- 63
Db      2  PPAGGSGSEKVVVRCRPLNGKEKADGRGRIVD---MDVDAGQVKVRNPKADASPPKA 57
QY      64  YQDFAYGERSQTDIYAGSVQPIRLHLEGNASVLAYGPTGAGKTHMLGSPGP--- 120
Db      58  FTFDQVYDNCQQRDVFDTARELIDSCIEGYNGTTFAYGQTGTGKSHTWEGKDEPELR 117
QY      121  GVTPRALMDLLO-LTREBAGRPWALSVMYSYLEIYQEKVLDLL-DPASGDVIREDCR 178
Db      118  GLIPNTFRVVEIIARDSGTK-----EFLVRSSYLEIYNEVRDLGDKSHKMKELKESPD 173
QY      179  GNILIFGLSQKPISSPADFERHFLPASRRTVCATRLNQSSSRSHAVLLVKVDQRELRAP 238
Db      174  RGVYVKDLQSFQCKNYEEMKVLLAGKDNQVQATLNMQDSSRSHSIFTITIECIKLES 233
QY      239  F-----RQREGKLYLIDLAGSEDRNRRTGNKGLRLKESGAINTSFLVLGKAVD 285

```

```

Db      234  AAAQPKGAKKDSNHRVVKGLNVLVDLAGSERQDKTGADRLKEGIKINLSLTALGNVIS 293
QY      286  ALNQGLP-RVPYRDSKLTLLQDSLGGSAHSILIANIAPERRPYLDTVSALNPAARSKV 344
Db      294  ALVDGKSGHIPYRDSKLTLLQDSLGKNTKTVMANIGPADWNYDETMSTLRVANRAKNI 353
QY      345  INRPFTNESLOPHALGPVKLSQKELLGPPPEAKRARGPEEEETGSPPEM-AAPASQKLS 403
Db      354  QNKPKINEDPKDAMLRFQOEETKKLKEQLAARAAGG-----GGPITMPSGGGSPTKI- 406
QY      404  PLQKLSMDP-----AMLERLLSLD-----RLLASQSG- 433
Db      407  -VERTEVDPDIDAIAKQMRLEAKMKSDISTEALDKAREEAAAKKQLQAIIDDQK 465
QY      434  -----APLSTPKRRMVLMTVVEKD-----LE-IER 460
Db      466  TEAQKAARDALKQAEARATAGATEKEQKAVLESRIKEMEGKIVVGVVNMLEKVD 525
QY      461  LKTKQKELEAKMLAQKAEKE 481
Db      526  LKQKSEDIKREAAIRKQEE 546

RESULT 12
H84777
probable kinesin-related cytokeleton protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84777
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1056 <STO>
A;Cross-references: UNIPROT:P82266; GB:AE002093; NID:g4510356; PIDN:AA21445.1; GSPDB:G
C;Genetics:
A;Gene: At2g36200
A;Map position: 2
C;Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match      22.7%; Score 561.5; DB 2; Length 1056;
Best Local Similarity 31.3%; Pred. No. 5.8e-30;
Matches 156; Conservative 86; Mismatches 152; Indels 105; Gaps 14;

QY      20  VRVAVRLRPVDGTAGASDPPCVRGMDSCSLSIANWRN---HQETLKQYQDFAYGERST 75
Db      13  VQVLLRCRPFSDDEL-RSNAPQVLTCDLQREAVSQNIAGHIDRV-FTFDKVFQPSAQ 70
QY      76  QQDIYAGSVQPIRLHLEGNASVLAYGPTGAGKTHMLGSG-----PEQPGVI 123
Db      71  QKDLVDQAVPIVNEVLEGFNCTIFAYGQTGTGKTYTMEGECRSKSPACGGLPAEAGVI 130
QY      124  PRALMDLLOLTREBAGRPWALSVMYSYLEIYQEKVLDLLD-----ASGDLV 172
Db      131  PRAVKQIF-----DTLEGQQAESVKVTFLELYNEBITDILLAPEDLSRVAABEKQKPLP 185
QY      173  IREDCGNLILGLSQKPISS-----PADFERHFLPASRRTVCATRLNQSSSRSHAVLLV 228
Db      186  LMEDGKGVLVRLGLEEIEIVTSANEIFTLLER-----GSSKRRTAETFLNKQSSSHLSFI 241
QY      229  KVDQRELRAPFR--REGKLYLIDLAGSEDRNRRTGNKGLRLKESGAINTSFLVLGKVDA 286
Db      242  THIKETATPEGELIKCGKLNVLVDLAGSENISSSGARDGAREAGEINKSLTLGRVISA 301
QY      287  LNOGLPRVPRYDSKLTLLQDSLGGSAHSILIANIAPERRPYLDTVSALNPAARSKVIN 346
Db      302  LVEHLGHVPRYDSKLTLLRLDSLGRTKTCIIATVSPAVHCLSETLSTLDYAHRAKNIRN 361

```



Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141, MUID:21016719, PMID:11130712  
A:Accession: D86151  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-885 <STO>  
A:Cross-references: UNIPROT:Q9LPC6; GB:AE005172; NID:98570446; PIDN:AAF76473.1; GSPDB:GN  
C:Genetics:  
A:Map position: 1

Query Match 22.7%; Score 560; DB 2; Length 885;  
Best Local Similarity 31.5%; Pred. No. 5.7e-30;  
Matches 169; Conservative 89; Mismatches 203; Indels 76; Gaps 15;  
Qy 3 RCRLSKIGATRRPPPARVRVAVRLRP--FVDGTAGASDPPCVR---GMDSCSLEIANWRN 57  
Db 44 RSSASIGADNGVPGVRVAVRLPRNADESADADFCVQLPELKLKLNW-- 101  
Qy 58 HQETLYQFDAYGERSTQDIYAGSVQPIRLHLLEGONASVLAYGPTGAGKTHM--LG 115  
Db 102 --DTETFEDEVLTAAASKQVVEVAKPVESVLEGYNGTVMAYGQTGKTFILORLG 159  
Qy 116 SPEQP--GVTPRALMDLLQUTREEGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGLVI 173  
Db 160 DETAARGIMVRSMEIIGTSLD-----TDSISVSYLYMETTQDLLDPTNDNIAI 212  
Qy 174 REOCR--GNILPGLSQKPISSFADFHFPLPASNRNRTVGATRLNQSRSHAVLLVKV-- 230  
Db 213 VEDPRTGDVSLPGATHVEIRNQNFLELLQLGETHRVAANTKLNTSSRSHAILMVHVKR 272  
Qy 231 -----DQERLAPFRQ-----REGKLYLIDLAGEDNRRRTGNKGLRLKESGAIN 274  
Db 273 SVVNEPFPVSNEMESSHFVRPSKPLVRSKLVLDLAGSERVHKSGEGHMLEAKSIN 332  
Qy 275 TSLFVLGKVDALNQGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSA 334  
Db 333 LSLSALGKICINATAENSHPVPLRDSKLTLLRDSFGGTARTSLIVTIGPSPRHGETTST 392  
Qy 335 LNFAARSKVINRPFNTNESLQPHALGP-----VKLSQKELLG-PPEAKARGPEEE 384  
Db 393 ILFGQRAKVENMLKIKEEDFYKSLKLEVLQDKVIAENERQLKAFDDDDVERINQAOQ 452  
Qy 385 EIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLSLDR-----LLASQ 430  
Db 453 RISEVEKNFAALEKEKLCQMEYMSVKLEKLIHQNHENKNGEVNGVVTASEF 512  
Qy 431 SQGAPLLSTPKRERMVLMKVEE-----KDLIERLKTQKELEAKMLAQKAEKE 481  
Db 513 TR-----LKESLENEMKLKRSABEVSKEGEDAGITRL---QKLEDEALQKKKLEEE 562

Search completed: November 5, 2004, 18:45:35  
Job time : 16.1982 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:16 ; Search time: 297.595 Seconds  
(without alignments)  
1813.185 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKLIAQAEKENHCPMT 487

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 33: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 34: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 35: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*
- 36: /cgn2\_6/ptodata/1/paa/US09 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	19	US-09-597-292-8
2	2472	100.0	487	33	US-10-797-893-8
3	2467	99.8	512	19	US-09-597-292-4
4	2467	99.8	512	33	US-10-797-893-4
5	2467	99.8	665	33	US-10-733-878-425
6	2309	93.4	482	27	US-10-170-2058-10738
7	2309	93.4	490	1	PCT-US02-41687-80
8	2309	93.4	490	1	PCT-US03-02234-80
9	2309	93.4	490	24	US-09-949-016-8784
10	2309	93.4	490	29	US-10-334-143-80
11	2309	93.4	490	35	US-10-940-774-8784
12	2235	90.4	559	21	US-09-724-676-90071
13	2235	90.4	559	21	US-09-724-676-90071
14	2235	90.4	588	21	US-09-724-676-90072
15	2235	90.4	597	21	US-09-724-676-90098
16	2235	90.4	597	21	US-09-724-676-90098
17	2235	90.4	597	21	US-09-724-676-90098
18	2235	90.4	626	21	US-09-724-676-90096
19	2235	90.4	626	21	US-09-724-676-90096
20	2235	90.4	627	21	US-09-724-676-90099
21	2235	90.4	656	21	US-09-724-676-90099
22	2235	90.4	656	21	US-09-724-676-90097
23	2235	90.4	656	21	US-09-724-676-90097
24	2141	86.6	460	1	PCT-US00-05988-1228
25	2141	86.6	460	24	US-09-925-300-1228
26	2128.5	86.1	627	21	US-09-724-676-90095
27	2128.5	86.1	627	21	US-09-724-676-90095
28	2081	84.2	414	27	US-10-170-2058-11425
29	2081	84.2	414	27	US-10-170-2058-11425
30	2014	81.5	665	22	US-09-791-537-149150
31	1768	71.5	346	19	US-09-597-292-6
32	1768	71.5	346	33	US-10-797-893-6
33	1763	71.3	370	19	US-09-597-292-2
34	1763	71.3	370	33	US-10-797-893-2
35	1438	58.2	651	22	US-09-791-537-149034
36	704	28.5	148	22	US-09-791-537-102346
37	687.5	27.8	147	36	US-60-182-902-537
38	673	27.2	134	36	US-60-163-233-3036
39	659	26.7	135	36	US-60-169-867-7228
40	634.5	25.7	543	30	US-10-438-246-16760
41	633	25.6	584	30	US-10-449-902-51666
42	625.5	25.3	377	27	US-10-155-881-27026
43	616	24.9	784	20	US-09-614-150A-40128
44	616	24.9	784	20	US-09-614-150A-40128
45	616	24.9	784	22	US-09-791-537-74691

# ALIGNMENTS

RESULT 1  
US-09-597-292-8  
; Sequence 8, Application US/09597292  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: US 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Human  
US-09-597-292-8

Query Match 100.0%; Score 2472; DB 19; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-204; Indels 0; Gaps 0;  
 Matches 487; Conservative 0; Mismatches 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 60

QY 61 TLKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSP 120  
 DB 61 TLKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSP 120

QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180  
 DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180

QY 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240  
 DB 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240

QY 241 QREGKLYLIDLAGSNDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGSNDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 300

QY 301 LTRLLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNESLOPHALG 360  
 DB 301 LTRLLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNESLOPHALG 360

QY 361 PVKLSOKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
 DB 361 PVKLSOKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420

QY 421 SLDRLASQSGQCAPLLSTPKRERMVLMKTVEEKDLEIERLTKKQLEAKMAKAEK 480  
 DB 421 SLDRLASQSGQCAPLLSTPKRERMVLMKTVEEKDLEIERLTKKQLEAKMAKAEK 480

QY 481 ENHCPTM 487  
 DB 481 ENHCPTM 487

RESULT 2  
 US-10-797-893-8  
 ; Sequence 8, Application US/10797893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10797,893  
 ; PRIOR FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 487  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-10-797-893-8

Query Match 100.0%; Score 2472; DB 33; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-204; Indels 0; Gaps 0;  
 Matches 487; Conservative 0; Mismatches 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 60

QY 61 TLKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSP 120

DB 61 TLKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSP 120  
 QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180  
 DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 180

QY 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240  
 DB 181 ILIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 240

QY 241 QREGKLYLIDLAGSNDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGSNDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 300

QY 301 LTRLLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNESLOPHALG 360  
 DB 301 LTRLLQDSIGGSAHSILIANIAPERRFYLDTVSALNFAARSKVINRPTNESLOPHALG 360

QY 361 PVKLSOKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
 DB 361 PVKLSOKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420

QY 421 SLDRLASQSGQCAPLLSTPKRERMVLMKTVEEKDLEIERLTKKQLEAKMAKAEK 480  
 DB 421 SLDRLASQSGQCAPLLSTPKRERMVLMKTVEEKDLEIERLTKKQLEAKMAKAEK 480

QY 481 ENHCPTM 487  
 DB 481 ENHCPTM 487

RESULT 3  
 US-09-597-292-4  
 ; Sequence 4, Application US/09597292  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/09/597,292  
 ; CURRENT FILING DATE: 2000-06-20  
 ; PRIOR APPLICATION NUMBER: US 09/295,612  
 ; PRIOR FILING DATE: 1999-04-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 512  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-597-292-4

Query Match 99.8%; Score 2467; DB 19; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-204; Indels 0; Gaps 0;  
 Matches 486; Conservative 0; Mismatches 0;

QY 2 GRCRLSKIGATRRPPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 61  
 DB 27 GRCRLSKIGATRRPPPARVAVRLRPFDVGTAGASDPPCVRGMDSCSLEIANWRNHQ 86

QY 62 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSP 121  
 DB 87 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAGPTGAGKTHMLGSP 146

QY 122 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 181  
 DB 147 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 206

QY 182 LIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 241  
 DB 207 LIPGLSOKPISSPADFERHFLPASRNRTVGATRLNQRSSRSHAVLLVKVDQERLAPFR 266

QY 242 REGKLYLIDLAGSNDNRRTGNKGLRKESGAINTSLFVLGKVVDALNOGLPRVPYRDSK 301

Db 267 REGKLYLDLAGSDNRRTGKGLRKEGSAINTSLFVLGKVDALNQGRLPRVPYRDSKL 326  
 Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 361  
 Db 327 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 386  
 Qy 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
 Db 387 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446  
 Qy 422 LDRLLASQSGAGAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKLEAKMLAKAEBKE 481  
 Db 447 LDRLLASQSGAGAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKLEAKMLAKAEBKE 506  
 Qy 482 NHCPTM 487  
 Db 507 NHCPTM 512

RESULT 4  
 US-10-797-893-4  
 ; Sequence 4, Application US/10797893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; PRIOR FILING DATE: 2004-03-09  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 512  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-10-797-893-4

Query Match 99.8%; Score 2467; DB 33; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-204;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHQT 61  
 Db 27 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHQT 86  
 Qy 62 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121  
 Db 87 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 146  
 Qy 122 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPAGSDLVIREDCRNI 181  
 Db 147 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPAGSDLVIREDCRNI 206  
 Qy 182 LIPGLSOKPTSSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPFRQ 241  
 Db 207 LIPGLSOKPTSSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPFRQ 266  
 Qy 242 REGKLYLDLAGSDNRRTGKGLRKEGSAINTSLFVLGKVDALNQGRLPRVPYRDSKL 301  
 Db 267 REGKLYLDLAGSDNRRTGKGLRKEGSAINTSLFVLGKVDALNQGRLPRVPYRDSKL 326  
 Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 361  
 Db 327 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 386  
 Qy 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
 Db 387 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446

Qy 422 LDRLLASQSGAGAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKLEAKMLAKAEBKE 481  
 Db 447 LDRLLASQSGAGAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKLEAKMLAKAEBKE 506  
 Qy 482 NHCPTM 487  
 Db 507 NHCPTM 512  
 RESULT 5  
 US-10-733-878-425  
 ; Sequence 425, Application US/10733878  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jean-Philippe Girard  
 ; APPLICANT: Francois Amalric  
 ; APPLICANT: Myriam Roussigne  
 ; APPLICANT: Thomas Clouaire  
 ; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR  
 ; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL  
 ; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION  
 ; FILE REFERENCE: BIOBANK 012A  
 ; CURRENT APPLICATION NUMBER: US/10/733,878  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 60/432699  
 ; PRIOR FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: 60/485027  
 ; PRIOR FILING DATE: 2003-07-03  
 ; NUMBER OF SEQ ID NOS: 535  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 425  
 ; LENGTH: 665  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-733-878-425

Query Match 99.8%; Score 2467; DB 33; Length 665;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-204;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHQT 61  
 Db 26 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEITANRNHQT 85  
 Qy 62 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121  
 Db 86 LKQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 145  
 Qy 122 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPAGSDLVIREDCRNI 181  
 Db 146 VIPRALMDLLQLTREAGRPWALSVTMSYLEIYQEKVLDLDDPAGSDLVIREDCRNI 205  
 Qy 182 LIPGLSOKPTSSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPFRQ 241  
 Db 206 LIPGLSOKPTSSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDORERLAPFRQ 265  
 Qy 242 REGKLYLDLAGSDNRRTGKGLRKEGSAINTSLFVLGKVDALNQGRLPRVPYRDSKL 301  
 Db 266 REGKLYLDLAGSDNRRTGKGLRKEGSAINTSLFVLGKVDALNQGRLPRVPYRDSKL 325  
 Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 361  
 Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDITVSALNFAARSKEVINRPTNESLOPHALGP 385  
 Qy 362 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
 Db 386 VKLSQKELLGPPPEAKRGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 445  
 Qy 422 LDRLLASQSGAGAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKLEAKMLAKAEBKE 481  
 Db 446 LDRLLASQSGAGAPLLSTPKRERVMVLMKTVEEKDLEIRLTKKLEAKMLAKAEBKE 505  
 Qy 482 NHCPTM 487

Db 506 NHCPTM 511

RESULT 6  
US-10-170-205E-10738  
; Sequence 10738, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 10738  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-10738

Query Match 93.4%; Score 2309; DB 27; Length 482;  
Best Local Similarity 100.0%; Pred. No. 1.8e-190;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
Db 26 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85

QY 62 LKQFDIFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHMLGSPQPG 121  
Db 86 LKQFDIFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHMLGSPQPG 145

QY 122 VIPRALMDLLQLTREGEAGRPWALSVTMSYLYIYQEKVLDDLDLPASGDLVIREDCRNI 181  
Db 146 VIPRALMDLLQLTREGEAGRPWALSVTMSYLYIYQEKVLDDLDLPASGDLVIREDCRNI 205

QY 182 LIPGLSQPISSFADFERRHFLPASRNRVTGATRLNQRSSSHAVLLVKVDORERLAPFRQ 241  
Db 206 LIPGLSQPISSFADFERRHFLPASRNRVTGATRLNQRSSSHAVLLVKVDORERLAPFRQ 265

QY 242 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPRYRDSKL 301  
Db 266 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPRYRDSKL 325

QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPFTHNESLOPHALGP 361  
Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPFTHNESLOPHALGP 385

QY 362 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
Db 386 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 445

QY 422 LDRLLASOGSOGAPLLSTPKERMVLMKTVEEKOL 456  
Db 446 LDRLLASOGSOGAPLLSTPKERMVLMKTVEEKOL 480

RESULT 7  
PCT-US02-41687-80  
; Sequence 80, Application PC/TUS0241687  
; GENERAL INFORMATION:  
; APPLICANT: SUGEN, INC.  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1544  
; CURRENT APPLICATION NUMBER: PCT/US02/41687  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: 60/343,169  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 80

; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-41687-80

Query Match 93.4%; Score 2309; DB 1; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.8e-190;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
Db 34 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93

QY 62 LKQFDIFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHMLGSPQPG 121  
Db 94 LKQFDIFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHMLGSPQPG 153

QY 122 VIPRALMDLLQLTREGEAGRPWALSVTMSYLYIYQEKVLDDLDLPASGDLVIREDCRNI 181  
Db 154 VIPRALMDLLQLTREGEAGRPWALSVTMSYLYIYQEKVLDDLDLPASGDLVIREDCRNI 213

QY 182 LIPGLSQPISSFADFERRHFLPASRNRVTGATRLNQRSSSHAVLLVKVDORERLAPFRQ 241  
Db 214 LIPGLSQPISSFADFERRHFLPASRNRVTGATRLNQRSSSHAVLLVKVDORERLAPFRQ 273

QY 242 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPRYRDSKL 301  
Db 274 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVVVDALNOGLPRVPRYRDSKL 333

QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPFTHNESLOPHALGP 361  
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPFTHNESLOPHALGP 393

QY 362 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
Db 394 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453

QY 422 LDRLLASOGSOGAPLLSTPKERMVLMKTVEEKOL 456  
Db 454 LDRLLASOGSOGAPLLSTPKERMVLMKTVEEKOL 488

RESULT 8  
PCT-US03-02234-80  
; Sequence 80, Application PC/TUS0302234  
; GENERAL INFORMATION:  
; APPLICANT: GRIGORIEV, IGOR V.  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1571  
; CURRENT APPLICATION NUMBER: PCT/US03/02234  
; CURRENT FILING DATE: 2003-01-28  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US03-02234-80

Query Match 93.4%; Score 2309; DB 1; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.8e-190;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
Db 34 GRCLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93

QY 62 LKQFDIFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHMLGSPQPG 121  
Db 94 LKQFDIFYGERSTQODIYAGSVQPIRLHLEQONASVLAAYGTGAGKTHMLGSPQPG 153



Qy 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVLDDLLDPASGDLVIREDCRNI 181  
Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVLDDLLDPASGDLVIREDCRNI 213  
Qy 182 LIPGLSOKPISSFADFERHFLPASRNTVGNTRLNQSSRSASHVLLVKVDORERLAPRQ 241  
Db 214 LIPGLSOKPISSFADFERHFLPASRNTVGNTRLNQSSRSASHVLLVKVDORERLAPRQ 273  
Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLDPASGDLVIREDCRNI 301  
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLDPASGDLVIREDCRNI 333  
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 361  
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 393  
Qy 362 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
Db 394 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453  
Qy 422 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 456  
Db 454 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 488

## RESULT 9

US-09-949-016-8784  
; Sequence 8784, Application US/09949016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8784  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8784

Query Match 93.4%; Score 2309; DB 24; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.8e-190;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLBIANWRNHQET 61  
Db 34 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLBIANWRNHQET 93  
Qy 62 LKYOQDAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHMLGSPRQ 121  
Db 94 LKYOQDAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHMLGSPRQ 153  
Qy 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVLDDLLDPASGDLVIREDCRNI 181  
Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVLDDLLDPASGDLVIREDCRNI 213  
Qy 182 LIPGLSOKPISSFADFERHFLPASRNTVGNTRLNQSSRSASHVLLVKVDORERLAPRQ 241  
Db 214 LIPGLSOKPISSFADFERHFLPASRNTVGNTRLNQSSRSASHVLLVKVDORERLAPRQ 273  
Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLDPASGDLVIREDCRNI 301  
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLDPASGDLVIREDCRNI 333  
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 361  
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 393  
Qy 362 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
Db 394 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453  
Qy 422 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 456  
Db 454 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 488

## RESULT 11

US-10-940-774-8784

Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 361  
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 393  
Qy 362 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
Db 394 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453  
Qy 422 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 456  
Db 454 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 488

## RESULT 10

US-10-334-143-80  
; Sequence 80, Application US/10334143  
; GENERAL INFORMATION:  
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1543  
; CURRENT APPLICATION NUMBER: US/10/334,143  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: 60/343,169  
; PRIOR FILING DATE: 2001-12-31  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 80  
; LENGTH: 490  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-334-143-80

Query Match 93.4%; Score 2309; DB 29; Length 490;  
Best Local Similarity 100.0%; Pred. No. 1.8e-190;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLBIANWRNHQET 61  
Db 34 GRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLBIANWRNHQET 93  
Qy 62 LKYOQDAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHMLGSPRQ 121  
Db 94 LKYOQDAFYGERSTQDDIYAGSVQPIRLHLEGNASVLAAYGPTGAGKTHMLGSPRQ 153  
Qy 122 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVLDDLLDPASGDLVIREDCRNI 181  
Db 154 VIPRALMDLLQLTREAGAGRPWALSVTMSYLEIQKVLDDLLDPASGDLVIREDCRNI 213  
Qy 182 LIPGLSOKPISSFADFERHFLPASRNTVGNTRLNQSSRSASHVLLVKVDORERLAPRQ 241  
Db 214 LIPGLSOKPISSFADFERHFLPASRNTVGNTRLNQSSRSASHVLLVKVDORERLAPRQ 273  
Qy 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLDPASGDLVIREDCRNI 301  
Db 274 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVDVLDLDPASGDLVIREDCRNI 333  
Qy 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 361  
Db 334 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKVFNINRPTNESLOPHALGP 393  
Qy 362 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 421  
Db 394 VKLSOKELGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSPLQKLSMDPAMLERLLS 453  
Qy 422 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 456  
Db 454 LDRLASQSGQAGAPLLSTPKRERVMKMTVEEKDL 488

Fri Nov 12 12:26:28 2004

us-10-797-893-8.rapm

```

; ORGANISM: Homo sapiens
; US-09-724-676-90071

Query Match      90.4%; Score 2235; DB 21; Length 559;
Best Local Similarity 99.8%; Pred. No. 5.9e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 45 MDSCSLEIANWRNHQETLKYQFDFAFYGERSTOODIYAGSVQPIRLHLLGQNASVLA YGP 104
DB 1 MDSCSLEIANWRNHQETLKYQFDFAFYGERSTOODIYAGSVQPIRLHLLGQNASVLA YGP 60

QY 105 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 164
DB 61 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 120

QY 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSHA 224
DB 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSHA 180

QY 225 VLLVKVDORERLAPFRQREGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGKV 284
DB 181 VLLVKVDORERLAPFRQREGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGKV 240

QY 285 DALNOGLPRVYRDSKLTLLQDSLGSASHSILLIANIAPERRFYLDTVSALNFAARSKEV 344
DB 241 DALNOGLPRVYRDSKLTLLQDSLGSASHSILLIANIAPERRFYLDTVSALNFAARSKEV 300

QY 345 INRPFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEIEIGSPPEMAAPASASQKLS 404
DB 301 INRPFTNESLOPHALGPVKLSQKELLGPPEAKRARGPEEIEIGSPPEMAAPASASQKLS 360

QY 405 LQKSSMDPAMLERLLSLDRLLASQSGAPLSTPKRERMVLMKTVEEKDLETERLTK 464
DB 361 LQKSSMDPAMLERLLSLDRLLASQSGAPLSTPKRERMVLMKTVEEKDLETERLTK 420

QY 465 QKELEAKMLAQKAEKENHCPTM 487
DB 421 QKELEAKMLAQKAEKENHCPTM 443

RESULT 13
US-09-724-676A-90071
; Sequence 90071, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90071
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676A-90071

Query Match      90.4%; Score 2235; DB 21; Length 559;
Best Local Similarity 99.8%; Pred. No. 5.9e-184;
Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 45 MDSCSLEIANWRNHQETLKYQFDFAFYGERSTOODIYAGSVQPIRLHLLGQNASVLA YGP 104
DB 1 MDSCSLEIANWRNHQETLKYQFDFAFYGERSTOODIYAGSVQPIRLHLLGQNASVLA YGP 60

QY 105 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 164
DB 61 TGAGKTHTMLGSPQPGVIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLL 120

QY 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSHA 224
DB 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSHA 180

; Sequence 8784, Application US/10940774
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774
; CURRENT FILING DATE: 2004-09-15
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 8784
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Human
; US-10-940-774-8784

Query Match      93.4%; Score 2309; DB 35; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.8e-190;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQET 61
DB 34 GRCLSKIGATRRPPPARVRVAVLRPFVDGTAGSDPPCVRGMDSCSLEIANWRNHQET 93

QY 62 LKYQFDFAFYGERSTOODIYAGSVQPIRLHLLGQNASVLA YGPTGAGKTHTMLGSPQPG 121
DB 94 LKYQFDFAFYGERSTOODIYAGSVQPIRLHLLGQNASVLA YGPTGAGKTHTMLGSPQPG 153

QY 122 VIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 181
DB 154 VIPRALMDLLQLTREBGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 213

QY 182 LIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSHA VLLVKVDORERLAPFRQ 241
DB 214 LIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSSHA VLLVKVDORERLAPFRQ 273

QY 242 REGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVYRDSKL 301
DB 274 REGKLYLIDLAGSEDRNRRTGNKGLRKESGAINTSFLVLGKVVDALNOGLPRVYRDSKL 333

QY 302 TRLLQDSLGSASHSILLIANIAPERRFYLDTVSALNFAARSKEVINRPFTNESLOPHALGP 361
DB 334 TRLLQDSLGSASHSILLIANIAPERRFYLDTVSALNFAARSKEVINRPFTNESLOPHALGP 393

QY 362 VKLSQKELLGPPEAKRARGPEEIEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 421
DB 394 VKLSQKELLGPPEAKRARGPEEIEIGSPPEMAAPASASQKLSPLQKSSMDPAMLERLLS 453

QY 422 LDRLLASQSGAPLSTPKRERMVLMKTVEEKDL 456
DB 454 LDRLLASQSGAPLSTPKRERMVLMKTVEEKDL 488

RESULT 12
US-09-724-676-90071
; Sequence 90071, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90071
; LENGTH: 559
; TYPE: PRT

```

Qy 225 VLLVKVQDQERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKV 284  
Db 181 VLLVKVQDQERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKV 240  
Qy 285 DALNOGLPRVPRYRDSKLTLLQDSIGGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344  
Db 241 DALNOGLPRVPRYRDSKLTLLQDSIGGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300  
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSP 404  
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSP 360  
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464  
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420  
Qy 465 QKELEAKMLAQKAEKENHCPTM 487  
Db 421 QKELEAKMLAQKAEKENHCPTM 443

## RESULT 14

US-09-724-676-90072

; Sequence 90072, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 90072

; LENGTH: 588

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-90072

Query Match 90.4%; Score 2235; DB 21; Length 588;

Best Local Similarity 99.8%; Pred. No. 6.4e-184;

Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 45 MDSCLSLEIANWRNHQETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGYP 104  
Db 1 MDSCLSLEIANWRNHQETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGYP 60  
Qy 105 TGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 164  
Db 61 TGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 120  
Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 224  
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 180  
Qy 225 VLLVKVQDQERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKV 284  
Db 181 VLLVKVQDQERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKV 240  
Qy 285 DALNOGLPRVPRYRDSKLTLLQDSIGGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344  
Db 241 DALNOGLPRVPRYRDSKLTLLQDSIGGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300  
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSP 404  
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSP 360  
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464  
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420  
Qy 465 QKELEAKMLAQKAEKENHCPTM 487  
Db 421 QKELEAKMLAQKAEKENHCPTM 443

## RESULT 15

US-09-724-676A-90072

; Sequence 90072, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 90072

; LENGTH: 588

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-90072

Query Match 90.4%; Score 2235; DB 21; Length 588;

Best Local Similarity 99.8%; Pred. No. 6.4e-184;

Matches 442; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 45 MDSCLSLEIANWRNHQETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGYP 104  
Db 1 MDSCLSLEIANWRNHQETLKYQFDIFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGYP 60  
Qy 105 TGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 164  
Db 61 TGAGKTHMLGSPQPGVIPRALMDLLQLTREEGAEGRPWALSVTMSYLEIYQEKVLDLL 120  
Qy 165 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 224  
Db 121 DPASGDLVIREDCRGNILIPGLSQKPISSPADFERHFLPASRNRRTVGATRLNQRSSSHA 180  
Qy 225 VLLVKVQDQERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKV 284  
Db 181 VLLVKVQDQERLAPPRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINSTSLFVLGKV 240  
Qy 285 DALNOGLPRVPRYRDSKLTLLQDSIGGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 344  
Db 241 DALNOGLPRVPRYRDSKLTLLQDSIGGSAHSILITANIAPERRFYLDTVSALNFAARSKEV 300  
Qy 345 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSP 404  
Db 301 INRPFTNESLQPHALGPVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASOKLSP 360  
Qy 405 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 464  
Db 361 LQKLSSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERMVLMKTVEEKDLEIERLTK 420  
Qy 465 QKELEAKMLAQKAEKENHCPTM 487  
Db 421 QKELEAKMLAQKAEKENHCPTM 443

Search completed: November 5, 2004, 19:04:21

Job time : 299.595 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 99.9557 Seconds  
(without alignments)  
1747.786 Million cell updates/sec

Title: US-10-797-893-8

Perfect score: 2472

Sequence: 1 MGRCLSKIGATRRPPPARV.....LEAKMLQAQAEKENHCPTM 487

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2472	100.0	487	7	ADC23344 Human kin
2	2472	100.0	487	8	ADQ60234 Human mic
3	2467	99.8	512	7	ADC23340 Human kin
4	2467	99.8	512	8	ADQ60230 Human mic
5	2467	99.8	665	8	ADQ09240 Human KNS
6	2309	93.4	490	7	ADK40973 Novel hum
7	2141	86.6	460	3	AAB56650 Human pro
8	1768	71.5	346	7	ADC23342 Human kin
9	1768	71.5	346	8	ADQ60232 Human mic
10	1763	71.3	370	7	ADC23338 Human kin
11	1763	71.3	370	8	ADQ60228 Human mic
12	616	24.9	784	4	ABB71112 Drosophil
13	593	24.0	677	4	ABB65183 Drosophil
14	578.5	23.4	1237	7	ADK40926 Novel hum
15	574.5	23.2	1232	7	ADK49938 Human lun
16	574.5	23.2	1232	7	ADK49938 Human lun
17	574.5	23.2	1232	7	ADJ18924 Human dis
18	574.5	23.2	1235	7	ADJ70347 Human hea
19	573.5	23.2	1234	5	ADK40901 Novel hea
20	572	23.1	522	5	ABG70993 Human HsK
21	568.5	23.0	473	5	ABG70991 Human HsK
22	568.5	23.0	898	4	ABG70992 Human tar
23	568	23.0	898	5	ABU53123 Intracell
24	568	23.0	898	5	AAU76957 Novel hum
25	568	23.0	898	7	ADC31082 Human nov
					Adi15915 Human PP

26	567.5	23.0	998	8	ADO44167	Ado44167 Structura
27	564.5	22.8	548	7	ADM04007	Adm04007 Human pro
28	564.5	22.8	905	5	AAU76967	Aau76967 Novel hum
29	561.5	22.7	1232	5	ABG70990	Abg70990 Human HsK
30	561	22.7	1388	5	AAE14400	Aae14400 Human kin
31	561	22.7	1388	5	AAU79590	Aau79590 Human kin
32	561	22.7	1388	6	ABR48222	AbR48222 Human bla
33	561	22.7	1388	7	ADB80468	AdB80468 Ovarian c
34	561	22.7	1388	7	ADC35116	AdC35116 Human pro
35	561	22.7	1388	8	ADL83290	AdL83290 Human BRO
36	561	22.7	1388	8	ADQ20128	AdQ20128 Human sof
37	561	22.7	1388	8	ADQ09226	AdQ09226 Human KNS
38	559.5	22.6	1401	7	ADJ94914	AdJ94914 Novel NOV
39	557.5	22.6	672	7	ADJ69858	AdJ69858 Human hea
40	557.5	22.6	1232	7	ADJ49937	AdJ49937 Human kin
41	557.5	22.5	1232	7	ADJ49932	AdJ49932 Human lun
42	557	22.5	1324	7	ADJ95086	AdJ95086 Novel NOV
43	556.5	22.5	1066	4	AAG67418	Aag67418 Amino aci
44	556.5	22.5	1066	6	ABG72693	Abg72693 Fruitfly
45	556.5	22.5	1066	7	ADG98862	Adg98862 Fruit fly

ALIGNMENTS

RESULT 1

ADC23344  
ID ADC23344 standard; protein; 487 AA.

XX AC ADC23344;

XX DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 8).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
XX KW cytosstatic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
XX KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX PN US6387644-B1.

XX PD 14-MAY-2002.

XX PF 28-NOV-2000; 2000US-00724224.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C;

XX WPI; 2003-706919/67.

XX N-PSDB; ADC23343.

XX PT Identifying a candidate agent as modulator of function of a target  
XX PT protein for treating cellular proliferation disorders by adding a  
XX PT candidate agent to a mixture of the target protein that  
XX PT directly/indirectly produces ADP or phosphate.

XX PS Claim 1; SEQ ID NO 8; 26pp; English.

XX CC This invention relates to a novel method for high throughput screening  
XX CC systems used to identify compounds for the treatment of cellular  
XX CC proliferation disorders. Specifically, it refers to candidate agents that  
XX CC are capable of modulating the activity of target proteins having motor  
XX CC domains, such that the target protein directly or indirectly produces ADP  
XX CC or phosphate. Furthermore, this activity can be determined using  
XX CC fluorescence or absorbance readouts. The present invention describes a  
XX CC method that identifies modulators of the target protein, which is a

CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,  
CC immunomodulators and antiinflammatory. Accordingly, through gene  
CC therapy, they can be used for the treatment of cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
CC polypeptide sequence is human KID protein (SeqID 8) of the invention.  
XX  
SQ Sequence 487 AA;

Query Match 100.0%; Score 2472; DB 7; Length 487;  
Best Local Similarity 100.0%; Pred. No. 6e-224;  
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
DB 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
QY 61 TLKYQDFYGERSTQODIYAGSVQPIILHLEGQNASVLYAGPTGAGKTHMLGSPROP 120  
DB 61 TLKYQDFYGERSTQODIYAGSVQPIILHLEGQNASVLYAGPTGAGKTHMLGSPROP 120  
QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
QY 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPR 240  
DB 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPR 240  
QY 241 QREGKLYLIDLAGSENRRTGNKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300  
DB 241 QREGKLYLIDLAGSENRRTGNKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300  
QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNLSLOPHALG 360  
DB 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNLSLOPHALG 360  
QY 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
DB 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
QY 421 SLDRLLASQSQGAPLLSTPKRRRWMLMKTVEEKLIEIRLTKQKELEAKMLAQKAEK 480  
DB 421 SLDRLLASQSQGAPLLSTPKRRRWMLMKTVEEKLIEIRLTKQKELEAKMLAQKAEK 480  
QY 481 ENHCPTM 487  
DB 481 ENHCPTM 487

RESULT 2  
ADQ60234  
ID ADQ60234 standard; protein; 487 AA.  
XX AC ADQ60234;  
XX DE 23-SEP-2004 (first entry)  
XX DE Human microtubule motor protein #4.  
XX Human; microtubule motor protein; cellular proliferation disorder;  
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
XX OS Homo sapiens.  
XX PN US6762043-B1.  
XX PD 13-JUL-2004.  
XX PF 06-MAR-2002; 2002US-00093317.  
XX PR 20-APR-1999; 99US-00295612.

PR 20-JUN-2000; 2000US-00597292.  
PR 28-NOV-2000; 2000US-00724224.  
XX (CYTO-) CYTOKINETICS INC.  
XX Beraud C;  
XX WPI; 2004-532491/51.  
DR N-ESDB; ADQ60233.  
XX New isolated microtubule motor protein, useful for screening modulators  
PT for treating cellular proliferation disorders such as cancer,  
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
PT inflammation.  
PS Claim 1; SEQ ID NO 8; 26pp; English.  
XX The invention relates to human microtubule motor proteins and the nucleic  
CC acids encoding them. The invention also relates to a method of screening  
CC for modulators of a motor protein which has microtubule stimulated ATPase  
CC activity, a method of testing for ATPase activity of microtubule motor  
CC proteins, methods to identify candidate agents that bind to a target  
CC protein or act as a modulator of the binding characteristics or  
CC biological activity of a target protein, modulators of the target  
CC protein, and methods of treating cellular proliferation disorders such as  
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
CC and inflammation, for treating disorders associated with kinesin-like DNA  
CC binding protein (KID) and for inhibiting KID. The sequences are used for  
CC screening for modulators of motor proteins useful for treating cellular  
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
CC hypertrophy, immune disorders and inflammation, for treating disorders  
CC associated with KID and for inhibiting KID and for treating autoimmune  
CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents a human microtubule motor protein  
CC of the invention.

Sequence 487 AA;  
Query Match 100.0%; Score 2472; DB 8; Length 487;  
Best Local Similarity 100.0%; Pred. No. 6e-224;  
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
DB 1 MGRCLSKIGATRRPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHOE 60  
QY 61 TLKYQDFYGERSTQODIYAGSVQPIILHLEGQNASVLYAGPTGAGKTHMLGSPROP 120  
DB 61 TLKYQDFYGERSTQODIYAGSVQPIILHLEGQNASVLYAGPTGAGKTHMLGSPROP 120  
QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
DB 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 180  
QY 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPR 240  
DB 181 ILIPGLSQPISSFADFERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDQERLAPPR 240  
QY 241 QREGKLYLIDLAGSENRRTGNKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300  
DB 241 QREGKLYLIDLAGSENRRTGNKGLKESGAINTSFLVLGKVVDALNOGLPRVPYRDSK 300  
QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNLSLOPHALG 360  
DB 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTSALNFAARKEVINRPTNLSLOPHALG 360  
QY 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
DB 361 PVKLSQKELLGPPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLL 420  
QY 421 SLDRLLASQSQGAPLLSTPKRRRWMLMKTVEEKLIEIRLTKQKELEAKMLAQKAEK 480

Db 421 SLDRLLASQSGGAPLLSTPKRERVMKTVKEDLETERLTKQKELEAKOLAQAABEK 480

QY 481 ENHCPTM 487

Db 481 ENHCPTM 487

RESULT 3

ADC23340

ID ADC23340 standard; protein; 512 AA.

XX

AC ADC23340;

XX

AC ADC23340;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human kinesin-like DNA binding protein (KID) (SeqID 4).

XX

KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;

KW cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy;

KW cancer; hyperplasia; restenosis; cellular proliferation disorder;

KW cardiac hypertrophy; immune disorder; inflammation.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 2 /note= "Encoded by CA"

FT

XX US6387644-B1.

XX

PD 14-MAY-2002.

XX

XX 28-NOV-2000; 2000US-00724224.

XX

PR 20-APR-1999; 99US-00295612.

PR

PR 20-JUN-2000; 2000US-00597292.

XX

PA (CYTO-) CYTOKINETICS INC.

XX

PI Beraud C;

XX

DR WPI; 2003-706919/67.

DR

DR N-PSDB; ADC23339.

XX

XX

PT Identifying a candidate agent as modulator of function of a target

PT protein for treating cellular proliferation disorders by adding a

PT candidate agent to a mixture of the target protein that

PT directly/indirectly produces ADP or phosphate.

XX

PS Claim 1; SEQ ID NO 4; 26pp; English.

XX

CC This invention relates to a novel method for high throughput screening

CC systems used to identify compounds for the treatment of cellular

CC proliferation disorders. Specifically, it refers to candidate agents that

CC are capable of modulating the activity of target proteins having motor

CC domains, such that the target protein directly or indirectly produces ADP

CC or phosphate. Furthermore, this activity can be determined using

CC fluorescence or absorbance readouts. The present invention describes a

CC method that identifies modulators of the target protein, which is a

CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiac,

CC immunomodulators and antiinflammatory. Accordingly, through gene

CC therapy, they can be used for the treatment of cancer, hyperplasia,

CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This

CC polypeptide sequence is human KID protein (SeqID 4) of the invention.

XX

SQ Sequence 512 AA;

Query Match 99.8%; Score 2467; DB 7; Length 512;

Best Local Similarity 100.0%; Pred. No. 1.9e-223;

Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCVRGMDSCSLEIANRNHQET 61

Db 27 GRCLSKIGATRRPPPARVAVRLRPFVDTAGASDPPCVRGMDSCSLEIANRNHQET 86

QY 62 LKYOFDIFYGERSTQDDIYAGSVQPIILRHLLLEGQNASVLAYGPTGAGKTHMTLGSPEQPG 121

Db 87 LKYOFDIFYGERSTQDDIYAGSVQPIILRHLLLEGQNASVLAYGPTGAGKTHMTLGSPEQPG 146

QY 122 VIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 181

Db 147 VIPRALMDLLQLTREEGAERPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 206

QY 182 LIPGLSQKPISSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDQERLAPPRQ 241

Db 207 LIPGLSQKPISSPADFERHFLPASRNTVGATRLNQRSSRSHAVLLVKVDQERLAPPRQ 266

QY 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDKL 301

Db 267 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVPYRDKL 326

QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARKEVINRFTNESLQPHALGP 361

Db 327 TRLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARKEVINRFTNESLQPHALGP 386

QY 362 VKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421

Db 387 VKLSQKELLGPPPEAKRARGPEEBEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446

QY 422 LDRLLASQSGGAPLLSTPKRERVMKTVKEDLETERLTKQKELEAKOLAQAABEK 481

Db 447 LDRLLASQSGGAPLLSTPKRERVMKTVKEDLETERLTKQKELEAKOLAQAABEK 506

QY 482 NHCPTM 487

Db 507 NHCPTM 512

RESULT 4

ADC60230

ID ADC60230 standard; protein; 512 AA.

XX

AC ADC60230;

XX

DT 23-SEP-2004 (first entry)

XX

DE Human microtubule motor protein #2.

XX

KW Human; microtubule motor protein; cellular proliferation disorder;

KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;

KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;

KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX

OS Homo sapiens.

XX

PN US6762043-B1.

XX

PD 13-JUL-2004.

XX

PF 06-MAR-2002; 2002US-00093317.

XX

PR 20-APR-1999; 99US-00295612.

PR

PR 20-JUN-2000; 2000US-00597292.

PR

PR 28-NOV-2000; 2000US-00724224.

XX

PA (CYTO-) CYTOKINETICS INC.

XX

PI Beraud C;

XX

DR WPI; 2004-532491/51.

XX

PT New isolated microtubule motor protein, useful for screening modulators

PT for treating cellular proliferation disorders such as cancer,

PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and

PT inflammation.

XX

Fri Nov 12 12:26:27 2004

us-10-797-893-8.rag

Claim 1; SEQ ID NO 4; 26pp; English.

PS The invention relates to human microtubule motor proteins and the nucleic  
XX acids encoding them. The invention also relates to a method of screening  
CC for modulators of a motor protein which has microtubule stimulated ATPase  
CC activity, a method of testing for ATPase activity of microtubule motor  
CC proteins, methods to identify candidate agents that bind to a target  
CC protein or act as a modulator of the binding characteristics or  
CC biological activity of a target protein, modulators of the target  
CC protein, and methods of treating cellular proliferation disorders such as  
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
CC and inflammation, for treating disorders associated with kinesin-like DNA  
CC binding protein (KID) and for inhibiting KID. The sequences are used for  
CC screening for modulators of motor proteins useful for treating cellular  
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
CC hypertrophy, immune disorders and inflammation, for treating disorders  
CC associated with KID and for inhibiting KID and for treating autoimmune  
CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents a human microtubule motor protein  
CC of the invention. Note: The specification states that this sequence is  
CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not  
XX appear to be the case.

SQ Sequence 512 AA;

Query Match 99.8%; Score 2467; DB 8; Length 512;  
Best Local Similarity 100.0%; Pred. No. 1.9e-223;  
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPARVRVAVLRPPFVDGTAGADPPCVRGMDSCSLIANWRNHQT 61  
DB 27 GRCLSKIGATRRPPARVRVAVLRPPFVDGTAGADPPCVRGMDSCSLIANWRNHQT 86  
QY 62 LKQFDYFGERSTQDDIYAGSVQPIRLHLLLEQONASVLAAYGTGAGKTHMLGSPQPG 121  
DB 87 LKQFDYFGERSTQDDIYAGSVQPIRLHLLLEQONASVLAAYGTGAGKTHMLGSPQPG 146  
QY 122 VIPALMDLLQLTREAGRGPWALSVTMSYLYEYQKVLDDLPASGDLVITREDCRGMI 181  
DB 147 VIPALMDLLQLTREAGRGPWALSVTMSYLYEYQKVLDDLPASGDLVITREDCRGMI 206  
QY 182 LIPGLSQKPTSSPFDHFLPASRNTVGTATRLNQRSSSHAVLLVKVDQRELAPFRQ 241  
DB 207 LIPGLSQKPTSSPFDHFLPASRNTVGTATRLNQRSSSHAVLLVKVDQRELAPFRQ 266  
QY 242 REGKLYLIDLAGSDNRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 301  
DB 267 REGKLYLIDLAGSDNRTGNKGLRKESGAINTSLFVLGKVVVDALNQGLPRVPYRDSKL 326  
QY 302 TRLLQDSLGSAHSILIANIAPERFVLDTVSALNFAARKEVINRPTNIESLOPHALGP 361  
DB 327 TRLLQDSLGSAHSILIANIAPERFVLDTVSALNFAARKEVINRPTNIESLOPHALGP 386  
QY 362 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421  
DB 387 VKLSQKELLGPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 446  
QY 422 LDRLLASQSGCAPLLSTPKRERMVLMKTVEEKDLEIRLTKQKLEAKMLAQABEKE 481  
DB 447 LDRLLASQSGCAPLLSTPKRERMVLMKTVEEKDLEIRLTKQKLEAKMLAQABEKE 506  
QY 482 NHCPTM 487  
DB 507 NHCPTM 512

RESULT 5  
ADQ09240  
ID ADQ09240 standard; protein; 665 AA.

XX AC ADQ09240;

XX

23-SEP-2004 (first entry)

Human KNSL4 protein SEQ ID NO:425.

XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;  
XX THAP responsive element; angiogenesis; inflammation; metastasis; cancer;  
XX apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;  
XX angiogenic; antiinflammatory; cardiovascular; cytostatic;  
XX neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;  
XX human.

XX Homo sapiens.

XX WO2004055050-A2.

XX 01-JUL-2004.

XX 10-DEC-2003; 2003WO-IB006434.

XX 10-DEC-2002; 2002US-0432699P.

XX 03-JUL-2003; 2003US-0485027P.

XX (ENDO-) ENDOCUBE SAS.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Girard J, Amalric F, Roussigne M, Clouaire T;

XX WPI; 2004-525034/50.

XX N-PSDB; ADQ09241.

XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)  
XX responsive gene for preventing or treating e.g. cancer or inflammation,  
XX comprises modulating the interaction of a THAP polypeptide with a nucleic  
XX acid.

XX Example 47; SEQ ID NO 425; 612pp; English.

XX The present invention describes a method for modulating the expression of  
XX a thanatos (death)-associated protein (THAP) responsive gene. The method  
XX comprises modulating the interaction of a THAP-family polypeptide or its  
XX biological fragment with a nucleic acid, and so enhancing or repressing  
XX the expression of the THAP responsive gene. Also described: (1) a method  
XX of modulating the expression of a gene responsive to a THAP/chemokine  
XX complex; (2) a pharmaceutical carrier; (3) a transcription factor decoy  
XX element essentially of a THAP responsive element; (4) a cell  
XX comprising a transcription factor decoy described above; (5) methods of  
XX modulating the interaction between a nucleic acid and a THAP-family  
XX polypeptide or its biological fragment, or a nucleic acid and a  
XX THAP/chemokine complex; (6) a vector packaging cell line comprising a  
XX cell comprising a viral vector which comprises a promoter operably linked  
XX to a nucleic acid encoding a THAP-family polypeptide or its biological  
XX fragment; (7) a method of constructing a cell which expresses a  
XX recombinant THAP-family polypeptide; (8) a method of ameliorating  
XX symptoms associated with a condition mediated by a THAP/chemokine complex  
XX; (9) methods of identifying a test compound that modulates transcription  
XX at a THAP responsive element or that modulates the transport of a  
XX chemokine into the nucleus; (10) methods for reducing the symptoms  
XX associated with a condition selected from excessive or insufficient  
XX angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
XX or insufficient apoptosis, cardiovascular disease and neurodegenerative  
XX diseases; symptoms associated with a condition resulting from the  
XX activity of a chemokine or a THAP-family polypeptide in an individual; or  
XX symptoms associated with transcriptional repression or activation  
XX mediated by a THAP-family polypeptide in an individual; (11) a vector  
XX comprising a THAP responsive promoter operably linked to a nucleic acid  
XX encoding a detectable product; (12) a genetically engineered cell  
XX comprising the vector described above or that expresses a THAP-family  
XX polypeptide or its biological fragment; (13) an in vitro transcription  
XX reaction comprising a nucleic acid comprising a THAP responsive promoter,  
XX ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-  
XX family polypeptide that does not bind to a chemokine. The pharmaceutical  
XX composition has antiangiogenic, antiinflammatory, cardiovascular,



CC cytotatic, neuroprotective and osteopathic activities, and can be used  
 CC as a THAP and THAP synthesis modulator. The composition can be used for  
 CC modulating the expression of a THAP responsive gene. Modulation is useful  
 CC for reducing symptoms of conditions such as excessive or insufficient  
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative  
 CC diseases. The present sequence is used in the exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 665 AA;

Query Match 99.8%; Score 2467; DB 8; Length 665;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-223;  
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPPVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 DB 26 GRCLSKIGATRRPPPARVRVAVLRPPVDTAGASDPPCVRGMDSCSLEIANWRNHQET 85  
 QY 62 LKYOQDAFYGRSTQODIYAGSVQPIRLHLEGGQNASVLAGPTGAGKTHMLGSPQPG 121  
 DB 86 LKYOQDAFYGRSTQODIYAGSVQPIRLHLEGGQNASVLAGPTGAGKTHMLGSPQPG 145  
 QY 122 VIPRALMDLLQLTREAGRPPWALSVTMSYLEIYQEKVLDLPPASGDLVIREDCRNI 181  
 DB 146 VIPRALMDLLQLTREAGRPPWALSVTMSYLEIYQEKVLDLPPASGDLVIREDCRNI 205  
 QY 182 LIPGLSKQPISSPADFERHFLPASNRRTVGATRLNQSRSSSHAVLLVKVDQERLAPFRQ 241  
 DB 206 LIPGLSKQPISSPADFERHFLPASNRRTVGATRLNQSRSSSHAVLLVKVDQERLAPFRQ 265  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 301  
 DB 266 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 325  
 QY 302 TRLLQDSLGGSAHSILANTAPERFVLDTSALNFAARKEVINRPTNIESLOPHALGP 361  
 DB 326 TRLLQDSLGGSAHSILANTAPERFVLDTSALNFAARKEVINRPTNIESLOPHALGP 385  
 QY 362 VKLSQKELLGPPEAKRARGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLS 421  
 DB 386 VKLSQKELLGPPEAKRARGPEEEIGSPPEMAAPASQKLSPLQKLSMDPAMLERLLS 445  
 QY 422 LDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDLEIRLTKQKELEAKMLAQKAEKE 481  
 DB 446 LDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDLEIRLTKQKELEAKMLAQKAEKE 505  
 QY 482 NHCPMT 487  
 DB 506 NHCPMT 511

RESULT 6  
 ADK40973  
 ID ADK40973 standard; protein; 490 AA.

XX ADK40973;  
 AC  
 DT 06-MAY-2004 (first entry)  
 XX Novel human kinase protein #80.  
 XX  
 KW cytotatic; immunomodulator; cardiant; neuroprotective; nootropic;  
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
 KW cancer; peripheral nervous system; central nervous system;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; viral infection; prion infection;  
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
 KW attention disorder; cognitive disorder; hypotension; hypertension;  
 KW psychotic disorder; neurological disorder; dyskinesia;  
 KW metabolic disorder; organ transplant rejection; enzyme.  
 XX

OS Homo sapiens.  
 XX WO2003057841-A2.  
 XX 17-JUL-2003.  
 PD 31-DEC-2002; 2002WO-US041687.  
 PF 31-DEC-2001; 2001US-0343169P.  
 PR (GRIG/) GRIGORIEV I V.  
 PA (SUDA/) SUDARSANAM S.  
 XX Grigoriev IV, Sudarsanam S;  
 XX WPI; 2003-587115/55.  
 DR New isolated, enriched or purified nucleic acid molecule encoding a  
 XX kinase polypeptide, useful for treating cancer, immune-related diseases,  
 PT cardiovascular disease, brain or neuronal-associated diseases and  
 PT metabolic disorders.  
 XX  
 PS Claim 1; SEQ ID NO 80; 491pp; English.  
 XX  
 CC The invention relates to novel isolated, enriched or purified nucleic acid  
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule  
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and  
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the  
 CC polypeptide in (a), except that it lacks one or more, but not all, of an  
 CC N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-  
 CC terminal domain, a coiled-coil structure region, a spacer region and a C-  
 CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,  
 CC polypeptides, methods and substance are useful for treating cancers,  
 CC immune-related diseases or disorders, cardiovascular disease, brain or  
 CC neuronal-associated diseases, and metabolic disorders. The disorders are  
 CC preferably cancers of the tissues or of hematopoietic origin, diseases of  
 CC the central or peripheral nervous system, Alzheimer's disease,  
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,  
 CC viral infections, infections caused by prions, infections caused by  
 CC bacteria, infections caused by fungi, ocular diseases, migraines, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognition  
 CC disorders, dyskinesias, hypertension, psychotic disorders, neurological  
 CC disorders, hypokinesias, metabolic disorders and organ transplant  
 CC rejection. This sequence corresponds to one of the kinase polypeptides of  
 CC the invention.  
 XX  
 SQ Sequence 490 AA;

Query Match 93.4%; Score 2309; DB 7; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-208;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVRVAVLRPPVDTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 DB 34 GRCLSKIGATRRPPPARVRVAVLRPPVDTAGASDPPCVRGMDSCSLEIANWRNHQET 93  
 QY 62 LKYOQDAFYGRSTQODIYAGSVQPIRLHLEGGQNASVLAGPTGAGKTHMLGSPQPG 121  
 DB 94 LKYOQDAFYGRSTQODIYAGSVQPIRLHLEGGQNASVLAGPTGAGKTHMLGSPQPG 153  
 QY 122 VIPRALMDLLQLTREAGRPPWALSVTMSYLEIYQEKVLDLPPASGDLVIREDCRNI 181  
 DB 154 VIPRALMDLLQLTREAGRPPWALSVTMSYLEIYQEKVLDLPPASGDLVIREDCRNI 213  
 QY 182 LIPGLSKQPISSPADFERHFLPASNRRTVGATRLNQSRSSSHAVLLVKVDQERLAPFRQ 241  
 DB 214 LIPGLSKQPISSPADFERHFLPASNRRTVGATRLNQSRSSSHAVLLVKVDQERLAPFRQ 273  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 301  
 DB 274 REGKLYLIDLAGSDNRRTGNKGLRLKESGAINTSFLVLGKVVDALNQGLPRVPRYRDKL 333

Fri Nov 12 12:26:27 2004

us-10-797-893-8.rag

QY	302	TELLQDSLGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPFNPTNESLOPHALGP	361
DB	334	TELLQDSLGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPFNPTNESLOPHALGP	393
QY	362	VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS	421
DB	394	VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS	453
QY	422	LDRLASQSQGAPLLSTPKRERVMVLMKTVEKDL	456
DB	454	LDRLASQSQGAPLLSTPKRERVMVLMKTVEKDL	488
RESULT 7			
AA	56650	standard; protein; 460 AA.	
XX	AA	56650;	
DT	13-MAR-2001	(first entry)	
XX	Human prostate cancer antigen protein sequence SEQ ID NO:1228.		
XX	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;		
KW	neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;		
KW	vulnerable; gastrointestinal; nephrotropic; antiinfective; gynaecological;		
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;		
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;		
KW	wound; infectious disease.		
XX	XX	Homo sapiens.	
OS	XX	WO200055174-A1.	
PN	XX	21-SEP-2000.	
PD	XX	08-MAR-2000; 2000WO-US005988.	
XX	XX	12-MAR-1999; 99US-0124270P.	
XX	XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	XX	(ROSE/) ROSEN C A.	
PI	XX	Rosen CA, Ruben SM;	
XX	XX	WPI; 2000-587513/55.	
DR	XX	N-PSDB; AAF15853.	
DR	XX	Prostate cancer associated gene sequences, referred to as prostate cancer	
PT	XX	antigens, useful for treatment, prevention, and diagnosis of disorders	
PT	XX	such as prostate cancer.	
XX	XX	Claim 11; Page 1649-1651; 2338pp; English.	
PS	XX	AAF15566 to AAF16505 encode the human prostate cancer associated	
CC	XX	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.	
CC	XX	The prostate cancer antigens can have neuroprotective, cytostatic,	
CC	XX	cardioactive, immunomodulatory, muscular, vulnerable, gastrointestinal,	
CC	XX	nephrotropic, antiinfective, gynaecological and antibacterial activities,	
CC	XX	and can be used in gene therapy. The prostate cancer antigen	
CC	XX	polynucleotides may be used for detection of prostate cancer, chromosome	
CC	XX	identification, as chromosome markers, and for numerous other diagnostic	
CC	XX	or research purposes. The prostate cancer antigens may be used to treat	
CC	XX	disorders such as neural, immune, muscular, reproductive,	
CC	XX	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative	
CC	XX	disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to	
CC	XX	AAB57303 represent sequences used in the exemplification of the present	
CC	XX	invention	
XX	XX	Sequence 460 AA;	
SQ	XX	Query Match 86.6%; Score 2141; DB 3; Length 460;	
	XX	Best Local Similarity 95.9%; Pred. No. 9.6e-193;	

Matches 422; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

QY 2 GRCLSKI GATRRPPPARVRVAVRLPFPVDTAGASDPFCVRGMDSCSLEIANWRNHQET 61

DB 4 GRCLSKI GATRRPPPARVRVAVRLPFPVDTAGASDPFCVRGMDSCSLEIANWRNHQET 63

QY 62 LKQFDAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAAYGTGAGKTHTMLGSPQPG 121

DB 64 LKQFDAFYGERSTOODIYAGSVQPIRLHLLLEGONASVLAAYGTGAGKTHTMLGSPQPG 123

QY 122 VIPRALMDLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181

DB 124 VIPRALMDLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 183

QY 182 LIPGLSQPISSPADPERHFLPASRNRVTGATLNQSSSHAVLKVVDQERLAPFRQ 241

DB 184 LIPGLSQPISSPADPERHFLPASRNRVTGATLNQSSSHAVLKVVDQERLAPFRQ 243

QY 242 REGKLYLIDLAGSEDNRTGNKGLRKESGAINTSFLVGLKVVDALNQGILPRVPRYRDKL 301

DB 244 REGKLYLIDLAGSEDNRTGNKGLRKESGAINTSFLVGLKVVDALNQGILPRVPRYRDKL 303

QY 302 TRLLQDSLGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPFNPTNESLOPHALGP 361

DB 304 TRLLQDSLGSAHSILIANIAPERFYLDTVSALNFAARSKEVINRPFNPTNESLOPHALGP 363

QY 362 VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLS 421

DB 364 VKLSQKELIGPPEAKRARGPEEEIGSPPEMAAPASASQKLSPLQKLSMDPAMLERLLQ 423

QY 422 LDRLASQSQGAPLLSTPK 441

DB 424 LGPSACLPGEPXGFSVEYFK 443

RESULT 8

ADC23342

ID ADC23342 standard; protein; 346 AA.

XX AC ADC23342;

XX DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 6).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;

XX KW cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;

XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;

XX KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX PN US6397644-B1.

XX PD 14-MAY-2002.

XX PF 28-NOV-2000; 2000US-00724224.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX XX (CYTO-) CYTOKINETICS INC.

XX PA Beraud C;

XX PI WPI; 2003-706919/67.

XX DR N-PSDB; ADC23341.

XX DR Identifying a candidate agent as modulator of function of a target

XX PT protein for treating cellular proliferation disorders by adding a

XX PT candidate agent to a mixture of the target protein that

XX PT directly/indirectly produces ADP or phosphate.

PS Claim 1; SEQ ID NO 6; 26pp; English.  
 XX This invention relates to a novel method for high throughput screening  
 CC systems used to identify compounds for the treatment of cellular  
 CC proliferation disorders. Specifically, it refers to candidate agents that  
 CC are capable of modulating the activity of target proteins having motor  
 CC domains, such that the target protein directly or indirectly produces ADP  
 CC or phosphate. Furthermore, this activity can be determined using  
 CC fluorescence or absorbance readouts. The present invention describes a  
 CC method that identifies modulators of the target protein, which is a  
 CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardants,  
 CC immunomodulators and antiinflammatories. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polypeptide sequence is human KID protein (SeqID 6) of the invention.  
 XX  
 SQ Sequence 346 AA;

Query Match 71.5%; Score 1768; DB 7; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 1e-157;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGSDPPCVRGMDSCSLEIANWRNHQ 60  
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGSDPPCVRGMDSCSLEIANWRNHQ 60  
 QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAAYGTGAGKTHMLGSPEQ 120  
 DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAAYGTGAGKTHMLGSPEQ 120  
 QY 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180  
 DB 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180  
 QY 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSRSHAVLLVKVDQERLAPFR 240  
 DB 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSRSHAVLLVKVDQERLAPFR 240  
 QY 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGILPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGILPRVPYRDSK 300  
 QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 9  
 ADQ60232  
 ID ADQ60232 standard; protein; 346 AA.  
 AC ADQ60232;  
 XX  
 DT 23-SEP-2004 (first entry)  
 XX Human microtubule motor protein #3.  
 DE  
 XX Human; microtubule motor protein; cellular proliferation disorder;  
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6762043-B1.  
 XX  
 PD 13-JUL-2004.  
 XX  
 XX 06-MAR-2002; 2002US-00093317.  
 XX  
 PR 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 PR 28-NOV-2000; 2000US-00724224.

XX (CYTO-) CYTOKINETICS INC.  
 XX Beraud C;  
 XX WPI; 2004-532491/51.  
 DR N-PSDB; ADQ60231.  
 XX  
 PT New isolated microtubule motor protein, useful for screening modulators  
 PT for treating cellular proliferation disorders such as cancer,  
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.  
 XX  
 PS Claim 1; SEQ ID NO 6; 26pp; English.  
 XX  
 CC The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase  
 CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention.  
 XX  
 SQ Sequence 346 AA;

Query Match 71.5%; Score 1768; DB 8; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 1e-157;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGSDPPCVRGMDSCSLEIANWRNHQ 60  
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDVGTAGSDPPCVRGMDSCSLEIANWRNHQ 60  
 QY 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAAYGTGAGKTHMLGSPEQ 120  
 DB 61 TLKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAAYGTGAGKTHMLGSPEQ 120  
 QY 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180  
 DB 121 GVIPRALMDLLQLTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN 180  
 QY 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSRSHAVLLVKVDQERLAPFR 240  
 DB 181 ILIPGLSQKPISSPADFERHFLPASRNRTVGATRLNQRSRSHAVLLVKVDQERLAPFR 240  
 QY 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGILPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLFVLGKVVDALNQGILPRVPYRDSK 300  
 QY 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 301 LTRLLQDSLGSSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 10  
 ADQ23338  
 ID ADQ23338 standard; protein; 370 AA.  
 XX  
 AC ADQ23338;  
 XX  
 DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 2).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;

XX KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;

XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;

XX KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX XX

XX FT Key .Location/Qualifiers

XX FT Misc-difference 1

XX FT /note= "Encoded by ATGCA"

XX XX

XX PN US6397644-B1.

XX XX

XX PD 14-MAY-2002.

XX XX

XX PF 28-NOV-2000; 2000US-00724224.

XX XX

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX XX

XX PA (CYTO-) CYTOKINETICS INC.

XX XX

XX PI Beraud C;

XX XX

XX DR WPI; 2003-706919/67.

XX DR N-PSDB; ADC23337.

XX XX

XX PT Identifying a candidate agent as modulator of function of a target

XX PT protein for treating cellular proliferation disorders by adding a

XX PT candidate agent to a mixture of the target protein that

XX PT directly/indirectly produces ADP or phosphate.

XX XX

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX XX

XX CC This invention relates to a novel method for high throughput screening

XX CC systems used to identify compounds for the treatment of cellular

XX CC proliferation disorders. Specifically, it refers to candidate agents that

XX CC are capable of modulating the activity of target proteins having motor

XX CC domains, such that the target protein directly or indirectly produces ADP

XX CC or phosphate. Furthermore, this activity can be determined using

XX CC fluorescence or absorbance readouts. The present invention describes a

XX CC method that identifies modulators of the target protein, which is a

XX CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiant,

XX CC immunomodulators and antiinflammatory. Accordingly, through gene

XX CC therapy, they can be used for the treatment of cancer, hyperplasias,

XX CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This

XX CC polypeptide sequence is human KID protein (SeqID 2) of the invention.

XX SQ

Query Match 71.3%; Score 1763; DB 7; Length 370;

Best Local Similarity 100.0%; Pred. No. 3.3e-157;

Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GCRLSKIGATRRPPPPARVAVRLRPVDTAGASDPPCVRGWDCSLEIANRNHOET 61

DB 26 GCRLSKIGATRRPPPPARVAVRLRPVDTAGASDPPCVRGWDCSLEIANRNHOET 85

QY 62 LKYQDFAVGERSTQDDIYAGSVQPIILHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 121

DB 86 LKYQDFAVGERSTQDDIYAGSVQPIILHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 145

QY 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLLDPASGDLVIREDCRGN 181

DB 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLLDPASGDLVIREDCRGN 205

QY 182 LIPGLSQPISSFADFERHFLPASRNRTVGATRLNQRSSRHVLLVKVDQREILAPFRQ 241

DB 206 LIPGLSQPISSFADFERHFLPASRNRTVGATRLNQRSSRHVLLVKVDQREILAPFRQ 265

QY 242 REKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVKGKVDALNQGILPRVPRYDSKL 301

DB 266 REKLYLIDLAGSDNRRTGNKGLRKESGAINTSFLVKGKVDALNQGILPRVPRYDSKL 325

QY 302 TRLLQSLGGSASISILIANIAPERRFYLDITVSALNFAARSKEVIN 346

DB 326 TRLLQSLGGSASISILIANIAPERRFYLDITVSALNFAARSKEVIN 370

RESULT 11

ADQ60228

ID ADQ60228 standard; protein; 370 AA.

XX AC ADQ60228;

XX DT 23-SEP-2004 (first entry)

XX XX Human microtubule motor protein #1.

XX DE Human; microtubule motor protein; cellular proliferation disorder;

XX KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;

XX KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;

XX KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX OS Homo sapiens.

XX XX

XX PN US6762043-B1.

XX XX

XX PD 13-JUL-2004.

XX XX

XX PF 06-MAR-2002; 2002US-00093317.

XX XX

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PR 28-NOV-2000; 2000US-00724224.

XX XX

XX PA (CYTO-) CYTOKINETICS INC.

XX XX

XX PI Beraud C;

XX XX

XX DR WPI; 2004-532491/51.

XX XX

XX PT New isolated microtubule motor protein, useful for screening modulators

XX PT for treating cellular proliferation disorders such as cancer,

XX PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and

XX PT inflammation.

XX XX

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX XX

XX CC The invention relates to human microtubule motor proteins and the nucleic

XX CC acids encoding them. The invention also relates to a method of screening

XX CC for modulators of a motor protein which has microtubule stimulated ATPase

XX CC activity, a method of testing for ATPase activity of microtubule motor

XX CC proteins, methods to identify candidate agents that bind to a target

XX CC protein or act as a modulator of the binding characteristics or

XX CC biological activity of a target protein, modulators of the target

XX CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders

XX CC and inflammation, for treating disorders associated with kinesin-like DNA

XX CC binding protein (KID) and for inhibiting KID. The sequences are used for

XX CC screening for modulators of motor proteins useful for treating cellular

XX CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac

XX CC hypertrophy, immune disorders and inflammation, for treating disorders

XX CC associated with KID and for inhibiting KID and for treating autoimmune

XX CC diseases, arthritis, graft rejection, inflammatory bowel disease and

XX CC proliferation induced after medical procedures including surgery and

XX CC angioplasty. This sequence represents a human microtubule motor protein

XX CC of the invention. Note: The specification states that this sequence is

XX CC encoded by the nucleic acid featured as SEQ ID NO:1, but this does not

XX CC appear to be the case.

XX XX

XX SQ Sequence 370 AA;

Query Match 71.3%; Score 1763; DB 8; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-157;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPARVAVRLRPVDTAGASDPPCVRGNDSCSLEIANRNHOET 61  
 DB 26 GRCLSKIGATRRPPARVAVRLRPVDTAGASDPPCVRGNDSCSLEIANRNHOET 85  
 QY 62 LKYQDFAFYGERSTQDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121  
 DB 86 LKYQDFAFYGERSTQDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 145  
 QY 122 VIPALMDLLQLTREAGRGPWALSVMTSMYLEIYQEKVLDLDPASGDLVIREDCRNI 181  
 DB 146 VIPALMDLLQLTREAGRGPWALSVMTSMYLEIYQEKVLDLDPASGDLVIREDCRNI 205  
 QY 182 LIPGLSKPISSPADFERHFLPASRNETVGATRLNQRSSRSHAVLLVKVQDRLAPFRQ 241  
 DB 206 LIPGLSKPISSPADFERHFLPASRNETVGATRLNQRSSRSHAVLLVKVQDRLAPFRQ 265  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLKESGAINTSFLVLGKVDALNOGLPRVPYRDSKL 301  
 DB 266 REGKLYLIDLAGSDNRRTGNKGLKESGAINTSFLVLGKVDALNOGLPRVPYRDSKL 325  
 QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

RESULT 12  
 ABB71112  
 ID ABB71112 standard; protein; 784 AA.  
 AC ABB71112;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 40128.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL15215.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 40128; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 784 AA;

Query Match 24.9%; Score 616; DB 4; Length 784;  
 Best Local Similarity 33.1%; Pred. No. 2.1e-48;  
 Matches 176; Conservative 84; Mismatches 163; Indels 88; Gaps 18;

QY 12 TRRP-----PPARVAVRLRPVDTAGASDPPCV-----RGMDSCSLEIANW--RN 57  
 DB 5 SRPFGTSSQTPNECVVCRPMNSRERSPEVNVYVNRGV----VELQVVDGN 60  
 QY 58 HQETLKYQDFAFYGERSTQDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHM--L 114  
 DB 61 KEQKQVFTYDAADASATQTLTLYHEVFPVPLVSSVLEGFNGCIFAYGOTGKTFTMEGVR 120  
 QY 115 GSPEQPGVPIPRALMDL-LQLTREAGRGPWALSVMTSMYLEIYQEKVLDLDPASGDLVI 173  
 DB 121 GNDELMLGIIIPRTFEQIWLHINRTEN-----FQFLVDVSYLEIYMEELRDLKPNKHLV 175  
 QY 174 REDCRGN-ILIPGLSKPISSPADFERHFLPASRNETVGATRLNQRSSRSHAVLLVKVDQ 232  
 DB 176 RE--RGSVYVYVNLHAINCKSVEDMIKVMQGNKRTVGTNNMHSRSHAFIMIKIEM 233  
 QY 233 RERLAPPROREGKLYLIDLAGSDNRRTGNKGLKESGAINTSFLVLGKVDALNOGLP 292  
 DB 234 CD-TEWNTIKVGNLIDLAGSDNRRTGNKGLKESGAINTSFLVLGKVDALNOGLP 292  
 QY 293 RVFYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNE 352  
 DB 293 HVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVINRPTNE 352  
 QY 353 SLQPHALGPVKLSQ-----KELGPPPEAKA-----ROPEEEI----- 386  
 DB 353 DPQ-----DAKLEYQBEIERLRLKLGIPQOQORSEKQVAKQKRVKPKETVTKEMSDS 407  
 QY 387 -----GSPEPMAAPASQKLSPLQKLSMDPAMLERLLSLRLLASQSSQ 433  
 DB 408 LQVSTIEQPVEDSDPE--GAESSEDKENAEVAKSNEELEERVEN-SKLAAKLAELE 463  
 QY 434 APLLSTPKRERMVLMKTVEKLEIER-----LTKQKELEAKMLAQKASE 479  
 DB 464 GLVVRGKKN-----LLDTYSERQIELEKLVIAERKKREIEIQOQLQEE 510

RESULT 13  
 ABB65183  
 ID ABB65183 standard; protein; 677 AA.  
 XX  
 AC ABB65183;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 22341.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX

PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL09286.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX Disclosure; SEQ ID NO 2341; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 677 AA;  
SQ  
Query Match 24.0%; Score 593; DB 4; Length 677;  
Best Local Similarity 33.9%; Pred. No. 2.5e-46;  
Matches 167; Conservative 78; Mismatches 190; Indels 58; Gaps 12;  
QY 20 VRVAVLRPFVDG--TAGADPPCVRGMD--SCSLEITANRNHQETLKYQDFAPYGERSTQ 76  
DB 21 VRVVTRTRPNKELSGALSAISVDKINRAITVMPNATANEPPKTYTFDVFDDGNSQ 80  
QY 77 QDIYAGSVQPIRLHLEGQNASVAYGPTGAGKTHMLGSPQEP---GVIPRALMDL--- 130  
DB 81 MDLYVDTARIVDKVLEGGYNTILAYGQTGKTGYTMSGNPDSPQTKGIIPNFAHIFGH 140  
QY 131 LQLTREBAGRPWALSVTWYSLEYIQEVLKLDLDPASG-DLVIREDCRGNILPLGLSQK 189  
DB 141 IAKAKEN-----QKFLRVSYMEIYNEEVRDLGKDVGKSLVKRDPDVGFKDLSGY 194  
QY 190 PISSEADFERHFLPASNRVVGATRLNQRSSRHAVLLVKVDQERLAPFRQ--REGKLY 247  
DB 195 VHVWADLENIMRLGNKRAVATKQWQESSRSHAFISITVERSELGEGDVQHVRMGKLQ 254  
QY 248 LIDLAGSEDRRTNGKRLKESGAINTSFLVIGKVVDALNOG-LPRVPRDSKLTLLQ 306  
DB 255 LVDLGASERQSKTQASGQRLKEATKINLSLVGLNVISALVDGKSTHPIYRNSKLTLLQ 314  
QY 307 DSLGGSASHILJANIPERRFYLDTVSALNFAARSKEVINRPTNE----SLQPHALGPV 362  
DB 315 DSLGGSNTVWCATISPADSNYMETIITLRYASRKNQIQRMINEEPKDALLRHFEQEI 374  
QY 363 KLSQKEL-LGPPEAKRARGPEEERIGSPPEMAAPASQKSLPQLKSSMDPAMLERLLS 421  
DB 375 ARLRKQLEEGDSLEEPSSEEDTADDELEAPLETELESSTTQAVEKPKKREK--- 431  
QY 422 LDRLLASQSGCAPLLSTPKRERVMKTVBEKDLERLKTQKLEAKM----- 472  
DB 432 -----TDAEKELAKRNEHQEIEHAKEQETLRNKLIVSLGKILV 473  
QY 473 ----LAQKAEKE 481  
DB 474 GGNLLEKQAQOE 486  
RESULT 14  
ADK40926  
ID ADK40926 standard; protein; 1237 AA.  
XX  
XX ADK40926;  
XX  
XX 06-MAY-2004 (first entry)  
XX

DE Novel human kinase protein #33.  
XX  
XX cytostatic; immunomodulator; cardiant; neuroprotective; nootropic;  
KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
KW cancer; peripheral nervous system; disease; central nervous system;  
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
KW amyotrophic lateral sclerosis; viral infection; prion infection;  
KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
KW attention disorder; cognition disorder; hypotension; hypertension;  
KW psychotic disorder; neurological disorder; dyskinesia;  
KW metabolic disorder; organ transplant rejection; enzyme.  
XX  
XX Homo sapiens.  
XX WO2003057841-A2.  
XX 17-JUL-2003.  
XX 31-DEC-2002; 2002WO-US041687.  
XX 31-DEC-2001; 2001US-0343169P.  
XX (GRIG/) GRIGORIEV I V.  
XX (SUDA/) SUDARSANAM S.  
XX Grigoriev IV, Sudarsanam S;  
XX WPI; 2003-587115/55.  
XX New isolated, enriched or purified nucleic acid molecule encoding a  
XX kinase polypeptide, useful for treating cancer, immune-related diseases,  
XX cardiovascular disease, brain or neuronal-associated diseases and  
XX metabolic disorders.  
XX Claim 1; SEQ ID NO 33; 491pp; English.  
XX The invention relates to novel isolated, enriched or purified nucleic acid  
XX molecules encoding a kinase polypeptide. The nucleic acid molecule  
XX comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
XX complement of (a); (c) hybridizes under stringent conditions to (a) and  
XX encodes a naturally occurring kinase polypeptide; (d) encodes the  
XX polypeptide in (a), except that it lacks one or more, but not all, of an  
XX N-terminal domain, C-terminal catalytic domain, a spacer region and a C-  
XX terminal tail; or (e) is a complement of (d). The nucleic acid molecules,  
XX polypeptides, methods and substance are useful for treating cancer,  
XX immune-related diseases or disorders, cardiovascular disease, brain or  
XX neuronal-associated diseases, and metabolic disorders. The disorders are  
XX preferably cancers of the tissues or of hematopoietic origin, diseases of  
XX the central or peripheral nervous system, Alzheimer's disease,  
XX Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,  
XX viral infections, infections caused by prions, infections caused by  
XX bacteria, infections caused by fungi, ocular diseases, migraines, pain,  
XX sexual dysfunction, mood disorders, attention disorders, cognition  
XX disorders, hypotension, hypertension, psychotic disorders, neurological  
XX disorders, dyskinesias, metabolic disorders and organ transplant  
XX rejection. This sequence corresponds to one of the kinase polypeptides of  
XX the invention.  
XX  
XX Sequence 1237 AA;  
Query Match 23.4%; Score 578.5; DB 7; Length 1237;  
Best Local Similarity 32.4%; Pred. No. 1.5e-44;  
Matches 157; Conservative 97; Mismatches 172; Indels 59; Gaps 13;  
QY 20 VRVAVLRPFVDGTAGASDPPC-----VRGWDSCSLEITANRNHQETLKYQDFAP 69  
DB 13 VRVALRCRPLVPKBISEGCQMLSFVPGTQVVVGTDK-----SFYDFV 57  
QY 70 YGERSTQDDIYAGSVQPIRLHLEGQNASVAYGPTGAGKTHMLGS-----PEQP--GV 122  
DB 58 FDPCTQEVEFNKAVAPLIKIFGKYNATVAYGQTGSGKTYSMGAYTAQEENPTVGI 117



Db 391 KNOQLVEENEKLSRGLSEAGQTAQML-----ERILLTEQANEKNVAKLEELRQHACKL 445

Qy 469 EAKMLAQKAEKE 481

Db : : : : :  
446 DLQKLVTLEDOE 458

Search completed: November 5, 2004, 18:44:42  
Job time : 101.956 secs



GenCore version 5.1.6.  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:25:02 ; Search time 3876.13 Seconds  
(without alignments)  
10482.189 Million cell updates/sec

Title: US-10-797-893-1  
Perfect score: 1115  
Sequence: 1 atgcagccgaggcggtctg.....ccaaggaggtgatcaattga 1115

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hc:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gsa1:\*
- 9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1107.4	99.3	2057	3	CR615781
2	1093.4	98.1	2025	3	CR590974
3	1045.4	93.8	1124	1	AL545712
4	1043.4	93.6	2016	3	CR590954
5	1002.8	89.9	1027	5	CR337406
6	980.8	88.0	1083	1	AL559772
7	978.6	87.8	1032	5	CR337353
8	942	84.5	959	5	CR337358
9	934	83.8	958	5	CR354166
10	921.4	82.6	999	1	AL559769
11	921.4	82.6	1094	5	CR376199
12	917	82.2	1031	4	BM910587
13	909	81.5	1019	4	BM546877
14	907.8	81.4	978	5	BM400806
15	905	81.2	1016	1	AL518361
16	889.8	79.8	1051	4	BM806996
17	889.6	79.8	980	5	BM404364
18	883.2	79.7	1041	4	BM478215
19	887.4	79.6	915	5	CR383181
20	862.4	77.3	926	1	AL578338
21	861.8	77.3	1040	4	BM546897
22	861.6	77.3	991	5	CR333336
23	851.8	76.4	865	1	AL555550
24	848.8	76.1	865	6	CD109092

25	844	75.7	990	5	CR399192
26	836.4	75.0	1139	4	BM553107
27	833.6	74.8	971	5	BO053671
28	833	74.7	838	5	CR341048
29	831.8	74.6	851	5	BUI88590
30	831.6	74.6	942	5	BU527647
31	823	73.8	995	1	AL559752
32	821.8	73.7	846	5	CR404325
33	819	73.5	921	5	BU860011
34	816.4	73.2	878	1	AL576578
35	815.2	73.1	872	5	BU521359
36	808.6	72.5	873	5	BQ955064
37	808.6	72.5	971	5	CR397059
38	801.2	71.9	880	4	BG683241
39	801	71.8	849	5	CR343522
40	789	70.8	961	5	BUI68881
41	787	70.6	813	4	B1835505
42	786.6	70.5	911	5	BU931772
43	786.4	70.5	956	1	AL558200
44	781.4	70.1	806	4	B1915916
45	781.4	70.1	913	5	CR445789

## ALIGNMENTS

RESULT 1  
LOCUS CR615781 2057 bp mRNA linear HTC 21-JUL-2004  
DEFINITION full-length cDNA clone CS0DK012YE23 of HeLa cells Cot 25-normalized of Homo sapiens (human).  
ACCESSION CR615781.1 GI:50496588  
VERSION CR615781.1  
KEYWORDS HTC; CNSLT; cDNA.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue  
REFERENCE 2 (bases 1 to 2057)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
FEATURES  
source location/Qualifiers  
1..2057  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK012YE23"  
/tissue\_type="HeLa cells Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

Query Match 99.3%; Score 1107.4; DB 3; Length 2057;  
Best Local Similarity 99.9%; Pred. No. 3.3e-294;  
Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	6	CCCGCGGCGCTCGACGACGACGACGACGACGATGGCGGAGCTTCAGCGCGCGG	65
Db	7	CCCGCGGCGCTCGACGACGACGACGACGACGATGGCGGAGCTTCAGCGCGCGG	66

```
Qy 66 ATCTCAGGAGTGTCGCTGTCGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 125
Db 67 ATCTCAGGAGTGTCGCTGTCGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 126
Qy 126 GCTCGCTAAGGTTGGCTGTCGCTGTCGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 185
Db 127 GCTCGCTAAGGTTGGCTGTCGCTGTCGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 186
Qy 186 GATCCCCCTGTCGCTGTCGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 245
Db 187 GATCCCCCTGTCGCTGTCGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 246
Qy 246 CACGAGAGACTCTCAAAATACGATTTGATGTCCTTCTATGGGAGAGGAGTACTCAGCAG 305
Db 247 CACGAGAGACTCTCAAAATACGATTTGATGTCCTTCTATGGGAGAGGAGTACTCAGCAG 306
Qy 306 GACATCTATGAGGTTGATGTCGAGCTTCTAAGGCACTTCTGTCGAGGAGGAGTACTCAGCAG 365
Db 307 GACATCTATGAGGTTGATGTCGAGCTTCTAAGGCACTTCTGTCGAGGAGGAGTACTCAGCAG 366
Qy 366 AGTGTGCTTGCCTATGAGGCTTGCAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 425
Db 367 AGTGTGCTTGCCTATGAGGCTTGCAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 426
Qy 426 GAGCAACTGTTGGTGTATCCGCGGCTCTATGAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 485
Db 427 GAGCAACTGTTGGTGTATCCGCGGCTCTATGAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 486
Qy 486 GGTGCGGAGGCGGCTTCTGTCGAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 545
Db 487 GGTGCGGAGGCGGCTTCTGTCGAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 546
Qy 546 GAGAAGGATATTAGACCTCTGAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 605
Db 547 GAGAAGGATATTAGACCTCTGAGGAGCTGGGAGAGCGCACAAATGCTGGGAGGAG 606
Qy 606 GGGGGGATATCTGATTCGGGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 665
Db 607 GGGGGGATATCTGATTCGGGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 666
Qy 666 GAGGCGACTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 725
Db 667 GAGGCGACTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 726
Qy 726 GGTCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 785
Db 727 GGTCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 786
Qy 786 CCATTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 845
Db 787 CCATTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846
Qy 846 CGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
Db 847 CGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
Qy 906 TTTGCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 965
Db 907 TTTGCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
Qy 966 GACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1025
Db 967 GACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026
Qy 1026 ATTGCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1085
Db 1027 ATTGCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086
Qy 1086 GCTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1114
Db 1087 GCTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1115
```

```
RESULT 2
CR590974
LOCUS
DEFINITION
2025 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DI009YJ01 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR590974
VERSION
HTC; CNSLT_CDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2025)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2025)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..2025
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI009YJ01"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 98.1%; Score 1093.4; DB 3; Length 2025;
Best Local Similarity 99.9%; Pred. No. 2.4e-290;
Matches 1094; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 20 GAGCGAGCAGAGCGGAGAGTGGCGGAGCTTTCAGCGGCGGCGGATCTCAGGAGCTGG 79
Db 1 GAGCGAGCAGAGCGGAGAGTGGCGGAGCTTTCAGCGGCGGCGGATCTCAGGAGCTGG 60
Qy 80 TCGCTGTCGGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCAGCTCCGCTAAGGGT 139
Db 61 TCGCTGTCGGCTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCAGCTCCGCTAAGGGT 120
Qy 140 GGCTGTGCGACTCGGCGCAATTTGTCGATGGAACAGCGGGAGCAAGTGATCCCCCTGTGT 199
Db 121 GGCTGTGCGACTCGGCGCAATTTGTCGATGGAACAGCGGGAGCAAGTGATCCCCCTGTGT 180
Qy 200 GCGGGGATGGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGGAGGAGTACTCAGGAGGAGCTCT 259
Db 181 GCGGGGATGGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGGAGGAGTACTCAGGAGGAGCTCT 240
Qy 260 CAAATACCACTTTGATGCGCTTCTATGGGAGGAGGAGTACTCAGCAGGAGACATCTATGCGG 319
Db 241 CAAATACCACTTTGATGCGCTTCTATGGGAGGAGGAGTACTCAGCAGGAGACATCTATGCGG 300
Qy 320 TTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGAGGAGTACTCAGGAGGAGTACTGCTTGCCTA 379
Db 301 TTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGAGGAGTACTCAGGAGGAGTACTGCTTGCCTA 360
Qy 380 TGGACCCACAGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 439
Db 361 TGGACCCACAGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
```

440 GATCCCGGGCTCTCATGACCTCTCTGAGCTCAACAGGAGAGGGTGCAGAGCGCG 499  
 421 GATCCCGGGCTCTCATGACCTCTCTGAGCTCAACAGGAGAGGGTGCAGAGCGCG 480  
 500 GCATGGGGCCCTTCTGTCACCATCTTACCTAGAGATCTACAGGAGAGGTATTAGA 559  
 481 GCATGGGGCCCTTCTGTCACCATCTTACCTAGAGATCTACAGGAGAGGTATTAGA 540  
 560 CCTCTCGGACCTCTCTGTCGAGAGCTCTGTAATCCGAGAGAGCTGCGGGGGAATATCT 619  
 541 CCTCTCGGACCTCTCTGTCGAGAGCTCTGTAATCCGAGAGAGCTGCGGGGGAATATCT 600  
 620 GATTCGGGTCTCTCCAGAGCCCATAGTACTTGTGTTGCTGTTGAGCGGCACTTCT 679  
 601 GATTCGGGTCTCTCCAGAGCCCATAGTACTTGTGTTGCTGTTGAGCGGCACTTCT 660  
 680 GCACGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGAGCGTCTCTCCGCG 739  
 661 GCACGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGAGCGTCTCTCCGCG 720  
 740 TCATGCTGTCTCTGCTCAAGTGTGAGCGGGAACGTTTGGCCCAATTTGCGCAGCG 799  
 721 TCATGCTGTCTCTGCTCAAGTGTGAGCGGGAACGTTTGGCCCAATTTGCGCAGCG 780  
 800 AGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGGACAAACCGGCGCACAGGCAA 859  
 781 AGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGGACAAACCGGCGCACAGGCAA 840  
 860 CAAGGGCTTCCGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTGCTGCGGCAA 919  
 841 CAAGGGCTTCCGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTGCTGCGGCAA 900  
 920 AGTGTGTAGTGTGCTCAAGTGTGAGCGGCTCTCTGCTGTACCTTATCGGACAGCACTCAC 979  
 901 AGTGTGTAGTGTGCTCAAGTGTGAGCGGCTCTCTGCTGTACCTTATCGGACAGCACTCAC 960  
 980 TCGCTATTGAGGAGCTCTCTGGGTGGGTGAGCCACAGTATCTTATGCAACATTCG 1039  
 961 TCGCTATTGAGGAGCTCTCTGGGTGGGTGAGCCACAGTATCTTATGCAACATTCG 1020  
 1040 CCTCAGAGAGCTCTTACCTAGACAGCTCTCGGACCTCAACTTGTGCTGCGAGTCAA 1099  
 1021 CCTCAGAGAGCTCTTACCTAGACAGCTCTCGGACCTCAACTTGTGCTGCGAGTCAA 1080  
 1100 GGAGGTGATCAATTG 1114  
 1081 GGAGGTGATCAATCG 1095

RESULT 3  
 LOCUS AL545712  
 DEFINITION AL545712 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01009YJ01 5-PRIME, mRNA sequence.  
 ACCESSION AL545712  
 VERSION AL545712.3 GI:45746191  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1124)  
 Li W.B., Gruber C., Jessee J. and Pollayes D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:31267547.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7766.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?e=CS01009YJ01&c=7766.r.  
 Location/Qualifiers  
 1..1124  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS01009YJ01"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 93.8%; Score 1045.4; DB 1; Length 1124;  
 Best Local Similarity 96.6%; Pred No. 3.8e-277; Indels 2; Gaps 2;  
 Matches 1059; Conservative 24; Mismatches 11;  
 QY 20 GACGACGAGGCGACGAGATGCGCGCAGCTTTCAGCGCGCGCATCTCAGGAGCTCG 79  
 DB 1 GACGACGAGGCGACGAGATGCGCGCAGCTTTCAGCGCGCGCATCTCAGGAGCTCG 60  
 QY 80 TCGCTGTGCTTAAAGAGATTTGGAGCTTCTGCTGCTCACCCTCAGCTCGGTAAAGGT 139  
 DB 61 TCGCTGTGCTTAAAGAGATTTGGAGCTTCTGCTGCTCACCCTCAGCTCGGTAAAGGT 120  
 QY 140 GCTGTGCTGCTGCGGCATTTGTTGGATGGAACACGCGGAGCAAGTATCCCCCTGTCT 199  
 DB 121 GCTGTGCTGCTGCGGCATTTGTTGGATGGAACACGCGGAGCAAGTATCCCCCTGTCT 180  
 QY 200 GCGGGGCATGACAGCTGCTCTCTAGAGATTTGCTAACTGGAGGAACCAACAGGAGACTCT 259  
 DB 181 GCGGGGCATGACAGCTGCTCTCTAGAGATTTGCTAACTGGAGGAACCAACAGGAGACTCT 240  
 QY 260 CAATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGACAG 319  
 DB 241 CAATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGACAG 300  
 QY 320 TTCACTGAGCCCATCTCAAGCACTTCTGTAAGGAGGAGATGCGAGTGTCTTGCCTTA 379  
 DB 301 TTCACTGAGCCCATCTCAAGCACTTCTGTAAGGAGGAGATGCGAGTGTCTTGCCTTA 360  
 QY 380 TGGACCCACAGGAGCTGGGAGAGCAGACAAATGCTGGGCGAGCCAGAGCAACCTGGGT 439  
 DB 361 TGGACCCACAGGAGCTGGGAGAGCAGACAAATGCTGGGCGAGCCAGAGCAACCTGGGT 420  
 QY 440 GATCCGCGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGAGGGTGCAGAGCGCG 499  
 DB 421 GATCCGCGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGAGGGTGCAGAGCGCG 480  
 QY 500 GCCATGGGCCCTTCTGTCACCATGCTTACTAGAGATCTACAGGAGAGGTATTAGA 559  
 DB 481 GCCATGGGCCCTTCTGTCACCATGCTTACTAGAGATCTACAGGAGAGGTATTAGA 540  
 QY 560 CCTCTCGGACCTCTCTGTCGAGAGCTCTGTAATCCGAGAGAGCTGCGGGGGAATATCT 619  
 DB 541 CCTCTCGGACCTCTCTGTCGAGAGCTCTGTAATCCGAGAGAGCTGCGGGGGAATATCT 600  
 QY 620 GATTCGGGTCTCTCCAGAGCCCATAGTACTTGTGTTGCTGTTGAGCGGCACTTCT 679  
 DB 601 GATTCGGGTCTCTCCAGAGCCCATAGTACTTGTGTTGCTGTTGAGCGGCACTTCT 660  
 QY 680 GCCAGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGAGCGTCTCTCCGCG 739  
 DB 661 GCCAGCCAGTCAAAATCGAGCTGTAGGAGCCACCGGCTCAACAGAGCGTCTCTCCGCG 720  
 QY 740 TCATGCTGTCTCTGCTCAAGTGTGAGCGGGAACGTTTGGCCCAATTTGCGCAGCG 799

```
Db 721 TCATGCTGTGCTCTGCTGAAGGKGGACCGCGGRACGTTTGGCCCAATTTGCCAGCG 780
Qy 800 AGAGGAAAACCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCACAGGCAA 859
Db 781 AGAGGAAAACCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCACAGGCAA 840
Qy 860 CAAGGGCTTCGGCT-AAAAGAGAGTGAGGACCAACACCTCCCTGTTTGTCTGGGCA 918
Db 841 CAAGGGCTTCGGSTAAAAGAGAGTGAGGACCAACACCTCCCTGTTTGTCTGGGCA 900
Qy 919 AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTGCTGTACCTTATCGGACAGCAAGCTCA 978
Db 901 AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTGCTGTACCTTATCGGRCAGCAAGCTCA 960
Qy 979 CTCGCTATTGACGACTCTGGGTGCTCAGCCGACAGTATCTTATTGCCAACATTG 1038
Db 961 CTCGCTATTGACGACTCTCTGGGKGSTCAG-CCACAGTATCTTATTGCCAACATTG 1019
Qy 1039 CCCCTGAGAGACGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTTGTGCCAGTCCA 1098
Db 1020 CCCCTGAGAGACGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTTGTGSCAGGKCCA 1079
Qy 1099 AGGAGGTGATCAATTG 1114
Db 1080 AGGGGKGRITCAATCG 1095

RESULT 4
CR590954
LOCUS
DEFINITION CR590954 2016 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0D1040Y106 of Placenta Cot 25-normalized
Of Homo sapiens (human).
ACCESSION CR590954
VERSION CR590954.1 GI:50471761
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2016)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization.
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
REFERENCE 2 (bases 1 to 2016)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..2016
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1040Y106"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 93.6%; Score 1043.4; DB 3; Length 2016;
Best Local Similarity 99.9%; Pred. No. 1.5e-276;
Matches 1044; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 70 CAGGAGCTGGTGGCTGAAGCAAGATTGGAGCTACTGTCGTCCACCTCCAGCTC 129
|||||
```

```
Db 1 CAGGAGCTGGTGGCTGTCGCTGAAGCAAGATTGGAGCTACTGTCGTCCACCTCCAGCTC 60
Qy 130 GCGTAAGGTTGGCTGTGCGACTCGGGCCATTGTTGATGAAACAGCGGGAGCAAGTGTATC 189
Db 61 GCGTAAGGTTGGCTGTGCGACTCGGGCCATTGTTGATGAAACAGCGGGAGCAAGTGTATC 120
Qy 190 CCCCTGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTTAACATGAGGAACCAACC 249
Db 121 CCCCTGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTTAACATGAGGAACCAACC 180
Qy 250 AGGAGACTCTCAAAATACCACTTTTATGTCCTTATGGGAGAGAGTACTCAGCAGACA 309
Db 181 AGGAGACTCTCAAAATACCACTTTTATGTCCTTATGGGAGAGAGTACTCAGCAGACA 240
Qy 310 TCTATGAGGTTTCACTGACGCCCATCTTAAGGCACTTCTGGAAGGCGAGATGCCAGTG 369
Db 241 TCTATGAGGTTTCACTGACGCCCATCTTAAGGCACTTCTGGAAGGCGAGATGCCAGTG 300
Qy 370 TGTCTGCTATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCAGAGC 429
Db 301 TGTCTGCTATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCAGAGC 360
Qy 430 AACCTGGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGACAGTCAACAAGGAGAGGGTG 489
Db 361 AACCTGGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGACAGTCAACAAGGAGAGGGTG 420
Qy 490 CCAGAGGCGCGGCATGCGGCCCTTCTGTACCAATGCTTACCTTAGAGATCTACCAGAGA 549
Db 421 CCAGAGGCGCGGCATGCGGCCCTTCTGTCTACCAATGCTTACCTTAGAGATCTACCAGAGA 480
Qy 550 AGGTATTAGACCTCTCTGACCCCTGCTTCGGGAGACCTGGTAATCCCGAAGAGACTGCCGGG 609
Db 481 AGGTATTAGACCTCTCTGACCCCTGCTTCGGGAGACCTGGTAATCCCGAAGAGACTGCCGGG 540
Qy 610 GGAATATCTGATTTCGGGGTCTCTCCAGAAGCCCATCAGTAGCTTTGCTGATTTTGAAC 669
Db 541 GGAATATCTGATTTCGGGGTCTCTCCAGAAGCCCATCAGTAGCTTTGCTGATTTTGAAC 600
Qy 670 GGCACTTCTCCAGCAGCTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGAGGCT 729
Db 601 GGCACTTCTCCAGCAGCTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGAGGCT 660
Qy 730 CTTCCCGCAGTCACTGCTGCTCTGCTCAAGGTGGACACGCGGGAACGTTTGGCCCCAT 789
Db 661 CTTCCCGCAGTCACTGCTGCTCTGCTCAAGGTGGACACGCGGGAACGTTTGGCCCCAT 720
Qy 790 TTCGCCAGCAGAGGGGAAAACCTCTA CTTGATTGACTTTGGGTGAGAGCAACCCGGC 849
Db 721 TTCGCCAGCAGAGGGGAAAACCTCTA CTTGATTGACTTTGGGTGAGAGCAACCCGGC 780
Qy 850 GCACAGGCAACAAGGGCTTCGGCTTAAAGAGATGGAGCCATCAACACCTCCCTGTTTG 909
Db 781 GCACAGGCAACAAGGGCTTCGGCTTAAAGAGATGGAGCCATCAACACCTCCCTGTTTG 840
Qy 910 TCCTGGCAAAAGTGTAGATGCTGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACA 969
Db 841 TCCTGGCAAAAGTGTAGATGCTGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACA 900
Qy 970 GCAAGCTCACTCGCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTG 1029
Db 901 GCAAGCTCACTCGCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTG 960
Qy 1030 CCAACATTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTCTGCTG 1089
Db 961 CCAACATTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTCTGCTG 1020
Qy 1090 CCAGGTCCAAGGAGGTGATCAATTG 1114
Db 1021 CCAGGTCCAAGGAGGTGATCAATCG 1045

RESULT 5
BX337406
```

LOCUS  
DEFINITION  
BX337406 1027 bp mRNA linear EST 07-APR-2004  
clone CS0D1041Y113 5-PRIME, mRNA sequence.  
ACCESSION  
BX337406  
VERSION  
BX337406.2 GI:46274820  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1027)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 2, 2003 this sequence version replaced gi:30341614.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
7766.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0D1041A070P1&c=7766.r.

# FEATURES

Location/Qualifiers  
1..1027  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1041Y113"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

# ORIGIN

Query Match 89.9%; Score 1002.8; DB 5; Length 1027;  
Best Local Similarity 99.5%; Pred. No. 2.1e-265;  
Matches 1023; Conservative 3; Mismatches 0; Indels 2; Gaps 2;  
QY 38 CGAGATGGCGGAGCTTACGCGCGGATCTCAGAGCTGTCGCTGCGGCTAAGCA 96  
DB 1 CGAGATGGCGGAGCTTACGCGCGGATCTCAGAGCTGTCGCTGCGGCTAAGCA 60  
QY 97 AGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGGCTGTCGCGACTGCGGC 156  
DB 61 AGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGGCTGTCGCGACTGCGGC 120  
QY 157 CATTGTGTGATGAACAGCGGAGAGTATCCCTGTCGCGGCGATGAGCAGCT 216  
DB 121 CATTGTGTGATGAACAGCGGAGAGTATCCCTGTCGCGGCGATGAGCAGCT 180  
QY 217 GCTCTCTAGAGATTGCTACTCGAGGAGACACAGAGACTCTCAATACCAAGTTGATG 276  
DB 181 GCTCTCTAGAGATTGCTACTCGAGGAGACACAGAGACTCTCAATACCAAGTTGATG 240  
QY 277 CTTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGGTTCAAGTCCGATCC 336  
DB 241 CTTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGGTTCAAGTCCGATCC 300  
QY 337 TAAGGCACTTGTGGAAGGCGAGATGCCAGTGTGTCCTATGAGACCAAGAGCTG 396  
DB 301 TAAGGCACTTGTGGAAGGCGAGATGCCAGTGTGTCCTATGAGACCAAGAGCTG 360  
QY 397 GGAAGACGACATGCTGGCAGCCAGACCACTGGGTGATCCCGGGGCTTCA 456  
DB 361 GGAAGACGACATGCTGGCAGCCAGACCACTGGGTGATCCCGGGGCTTCA 420

QY 457 TGGACCTCTCGAGCTCAACAGGAGAGGAGGTCGCCGAGGCGCGCCATGGGCGCTTTCTG 516  
DB 421 TGGACCTCTCGAGCTCAACAGGAGAGGAGGTCGCCGAGGCGCGCCATGGG-CCTTTCTG 479  
QY 517 TCACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTCTCTGGACCTGCTT 576  
DB 480 TCACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTCTCTGGACCTGCTT 539  
QY 577 CGGAGAGCTGTGTAATCCGAGAGACTGCGGGGGAATATCCTGATTCCGGGCTCTCTCC 636  
DB 540 CGGAGAGCTGTGTAATCCGAGAGACTGCGGGGGAATATCCTGATTCCGGGCTCTCTCC 599  
QY 637 AGAAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTCTCCGAGCAGTCTGAAATC 696  
DB 600 AGAAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTCTCCGAGCAGTCTGAAATC 659  
QY 697 GGAAGTGTAGAGCCACCCGGCTCAACAGGAGCTCTCCGCGAGTCAATGCTGCTCTCTGG 756  
DB 660 GGAAGTGTAGAGCCACCCGGCTCAACAGGAGCTCTCCGCGAGTCAATGCTGCTCTCTGG 719  
QY 757 TCAAGGTGAGCAGCGGGAACGTTTGGCGGCACTTCTCCGAGCAGGAGGAAACTCTACC 816  
DB 720 TCAAGGTGAGCAGCGGGAACGTTTGGCGGCACTTCTCCGAGCAGGAGGAAACTCTACC 779  
QY 817 TGATTGACTTGGCTGAGGAGCAACCGCGGCAACAGGAGGCGCTTTCGGCTAA 876  
DB 780 TGATTGACTTGGCTGAGGAGCAACCGCGGCAACAGGAGGCGCTTTCGGCTAA 839  
QY 877 AAGAGAGTGGAGCCATCAACACTCTCTGTTTGTCTGGGCAAAAGTGTAGATGCGCTGA 936  
DB 840 AAGAGAGTGGAGCCATCAACACTCTCTGTTTGTCTGGGCAAAAGTGTAGATGCGCTGA 899  
QY 937 ATCAGGCGCTCTCTGTTGAGCTTATCGGAGCAGCAAGCTCACTCGCCTATTGCGAGCT 996  
DB 900 ATCAGGCGCTCTCTGTTGAGCTTATCGGAGCAGCAAGCTCACTCGCCTATTGCGAGCT 959  
QY 997 CTCTGGGTGGCTCAGCCCACTATCTTATTCGCAAACTTGCCTCGAGAGACGCTTCT 1056  
DB 960 CTCTGGGTGGCTCAGCCCACTATCTTATTCGCAAACTTGCCTCGAGAGCGCTTCT 1019  
QY 1057 ACCTAGAC 1064  
DB 1020 ACCTAGAC 1027  
RESULT 6  
AL559772  
LOCUS  
DEFINITION  
AL559772 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens CDNA  
clone CS0D005Y111 5-PRIME, mRNA sequence.  
ACCESSION  
AL559772  
VERSION  
AL559772.3 GI:46185144  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1083)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:31283903.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 7766.r  
For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0D005CF06QPl&c=7766.r.

FEATURES  
Location/Qualifiers  
1..1083  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="CS0D005YL11"  
/tissue\_type="B CELLS (RAMOS CELL LINE)"  
/cell\_line="RAMOS CELL LINE"  
/clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"  
/note="vector: pcwvSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pcwvSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 88.0%; Score 980.8; DB 1; Length 1083;  
Best Local Similarity 96.4%; Pred. No. 2.5e-259;  
Matches 1010; Conservative 13; Mismatches 21; Indels 4; Gaps 2;

QY 10 CGGGGGCTCAGCAGCAGAGGCGAGCGCGAGATGGCGGAGCTTACAGCGCGCGGATCT 69  
DB 1 CGGGGGCTCAGCAGCAGAGGCGAGCGCGAGATGGCGGAGCTTACAGCGCGCGGATCT 60

QY 70 CAGGAGCTGGTGGCTGTCGGCTAAGCAAGATTGAGCTACTCTGCTCCACTCCAGCTC 129  
DB 61 CAGGAGCTGGTGGCTGTCGGCTAAGCAAGATTGAGCTACTCTGCTCCACTCCAGCTC 120

QY 130 CGTAAAGGCTGGTGGCTGTCGGCTAAGCAAGATTGAGCTACTCTGCTCCACTCCAGCTC 189  
DB 121 CGTAAAGGCTGGTGGCTGTCGGCTAAGCAAGATTGAGCTACTCTGCTCCACTCCAGCTC 180

QY 190 CCCCTGTGTGGGGGCTGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACC 249  
DB 181 CCCCTGTGTGGGGGCTGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACC 240

QY 250 AGGAGACTCTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACA 309  
DB 241 AGGAGACTCTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACA 300

QY 310 TCTATGCAAGTTCAGTGCAGGCGCATCTTAAGGCACTCTGCTGAGAGGCGAGATGCCAGTG 369  
DB 301 TCTATGCAAGTTCAGTGCAGGCGCATCTTAAGGCACTCTGCTGAGAGGCGAGATGCCAGTG 360

QY 370 TGCCTGCTTATGGACCCACAGAGCTGGGAGAGCCACACAATGCTGGGAGCCGAGAGC 429  
DB 361 TGCCTGCTTATGGACCCACAGAGCTGGGAGAGCCACACAATGCTGGGAGCCGAGAGC 420

QY 430 AACCTGGGGTATCCCGGGGCTCTCATGGACCTCTGAGCTCACAGGAGGAGGGTG 489  
DB 421 AACCTGGGGTATCCCGGGGCTCTCATGGACCTCTGAGCTCACAGGAGGAGGGTG 480

QY 490 CCGAGGCGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGA 549  
DB 481 CCGAGGCGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGA 540

QY 550 AGGTATTAGACCTCTCTGACCTCTGCTGGGAGACTGGTAATCCGAGAGACTCCCGGG 609  
DB 541 AGGTATTAGACCTCTCTGACCTCTGCTGGGAGACTGGTAATCCGAGAGACTCCCGGG 600

QY 610 GGAATATCTGATTCGGGCTCTCCAGAGCCCATCAGTACTGCTGATTTTGGAGC 669  
DB 601 GGAATATCTGATTCGGGCTCTCCAGAGCCCATCAGTACTGCTGATTTTGGAGC 660

QY 670 GGCACCTCTCTGCGAGCCAGTTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCT 729  
DB 661 GGCACCTCTCTGCGAGCCAGTTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCT 720

QY 730 CTTCCCGCAGTATGCTGCTCTGCTCAAGTGGAGCCAGGAGACTGTTGGCCCAT 789  
DB 721 CTTCCCGCAGTATGCTGCTCTGCTCAAGTGGAGCCAGGAGACTGTTGGCCCAT 780

790 TTCGCCAGCAGAGGAGAACTTACCTGATTGACTTGGCTGGTTCAGAGGACAAACCGGC 849  
781 TTCGCCAGCAGAGGAGAACTTACCTGATTGACTTGGCTGGTTCAGAGGACAAACCGGC 840

QY 850 GCACAGGCAACAAAGGCGCTTTCGGCTAAAGAGAGATGGAGCCATCAACACCTCCCTGTTTG 909  
DB 841 GCACAGGCAACAAAGGCGCTTTCGGCTAAAGAGAGATGGAGCCATCAACACCTCCCTGTTTG 900

QY 910 TCTGGGCAAGTGGTAGTGCCTGATCATCAGGCGCTCCCTGCTGCTTATCCGGACA 969  
DB 901 TCTGGGCAAGTGGTAGTGCCTGATCATCAGGCGCTCCCTGCTGCTTATCCGGACA 960

QY 970 GCAAGCTCACTCGCTTATGAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTTG 1029  
DB 961 GCAAGCTCACTCGCTTATGAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTTG 1016

QY 1030 CCAACATTGCGCTCAGAGACGCTTCTA 1057  
DB 1017 CAAMATTGCCCTKARARAGCTTTHCTA 1044

RESULT 7  
BX337353  
LOCUS  
DEFINITION  
BX337353 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
BX337353  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 1032)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 2, 2003 this sequence version replaced gi:30337600.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pcwvSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
7766.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0D1040B03QPl&c=7766.r.

FEATURES  
Location/Qualifiers  
1..1032  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="CS0D1040Y106"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoRV  
sites of the pcwvSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 87.8%; Score 978.6; DB 5; Length 1032;  
Best Local Similarity 98.6%; Pred. No. 1e-258;  
Matches 1003; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

QY 70 CAGGAGCTGGTGGCTGCTGCTAAGCAAGATTGAGTACTCTGCTCCACTCCAGCTC 129  
DB 1 CAGGAGCTGGTGGCTGCTGCTAAGCAAGATTGAGTACTCTGCTCCACTCCAGCTC 60

QY 130 CGTAAAGGCTGGTGGCTGCTGCTAAGCAAGATTGAGTACTCTGCTCCACTCCAGCTC 189

```
Db 61 GCCTAAGGTTGGCTGTGCGACTGCGCCATTTGTGATGGAACAGCGGGAGCAAGTGATC 120
Qy 190 CCCCTCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACCCACC 249
Db 121 CCCCTCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACCCACC 180
Qy 250 AGGAGACTCTCAAATACCAAGTTTGATGCGCTTCTATGGGAGAGAGTACTCAGCAGACA 309
Db 181 AGGAGACTCTCAAATACCAAGTTTGATGCGCTTCTATGGGAGAGAGTACTCAGCAGACA 240
Qy 310 TCTATCAGGTTTCAGTGCAGCCCATCTTAAGGACATTTGCTGGAAGGCGAATGCGCAGTG 369
Db 241 TCTATCAGGTTTCAGTGCAGCCCATCTTAAGGACATTTGCTGGAAGGCGAATGCGCAGTG 300
Qy 370 TGCTTCCTATGACCCACAGAGAGCTGGGAAGACGACACAATGCTGGGAGAGAGAGCCAGAGC 429
Db 301 TGCTTCCTATGACCCACAGAGAGCTGGGAAGACGACACAATGCTGGGAGAGAGAGCCAGAGC 360
Qy 430 AACCTGGGTTGATCCCGCGGCTCTCATGACCTCTCGAGCTCAACAGGAGAGAGGTTG 489
Db 361 AACCTGGGTTGATCCCGCGGCTCTCATGACCTCTCGAGCTCAACAGGAGAGAGGTTG 420
Qy 490 CCGAGGCGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGA 549
Db 421 CCGAGGCGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGA 480
Qy 550 AGGTATTAGACCTCTCTGACCTGCTTCCGGAGACCTGTAATCCGAGAAGACTGCCGG 609
Db 481 AGGTATTAGACCTCTCTGACCTGCTTCCGGAGACCTGTAATCCGAGAAGACTGCCGG 540
Qy 610 GGAATATCTGATTCGGGTCTCTCCAGAAGCCATCAGTACTGCTTGTGATTTTGGC 669
Db 541 GGAATATCTGATTCGGGTCTCTCCAGAAGCCATCAGTACTGCTTGTGATTTTGGC 600
Qy 670 GGCACCTCTCTGACGAGCTGAATTCGACTGTAGAGCCACCGGCTCAACAGCGCT 729
Db 601 GGCACCTCTCTGACGAGCTGAATTCGACTGTAGAGCCACCGGCTCAACAGCGCT 660
Qy 730 CTTCCCGCAGTCACTGCTGCTCTGCTCAAGTGTGACGAGCGGGAACGTTTGGCCCCAT 789
Db 661 CTTCCCGCAGTCACTGCTGCTCTGCTCAAGTGTGACGAGCGGGAACGTTTGGCCCCAT 720
Qy 790 TTGCGCAGCAGAGGAGAACTCTACCTGATGACTTTGGCTGGGTGAGAGCAACCCGC 849
Db 721 TTGCGCAGCAGAGGAGAACTCTACCTGATGACTTTGGCTGGGTGAGAGCAACCCGC 780
Qy 850 GCACAGGCAACAGGCGCTTCCGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTG 909
Db 781 GCACAGGCAACAGGCGCTTCCGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTG 840
Qy 910 TCTCGGCAAGTGTAGATGCTGTAATCAGGCGCTCCCTGCTGTACCTTATCGGAC 968
Db 841 TCTCGGCAAGTGTAGATGCTGTAATCAGGCGCTCCCTGCTGTACCTTATCGGAC 900
Qy 969 AGCAAGCTCACTCGCTATTGAGGACTCTC-TGGGTGGCTCAGCGCCACAGTATCCTTAT 1027
Db 901 AGCAAGCTCACTCGCTATTGAGGACTCTC-TGGGTGGCTCAGCGCCACAGTATCCTTAT 960
Qy 1028 TGCCAAATTTGCCCTTGAGAGAGCTTCTTACCTAGACAGTCTCCGACCTCAACT 1084
Db 961 TGCCAAATTTKSCCTTGAGAGAGCTTCTTATCTAGACAGTCTCGSCACTCACTTT 1017

RESULT 8
BX363358
LOCUS BX363358 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION BX363358 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION BX363358
VERSION BX363358.2 GI:46287820
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

## ORGANISM

Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1. (bases 1 to 959)

## AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

## TITLE

Full-length cDNA libraries and normalization

## JOURNAL

Unpublished (2001)

## COMMENT

On May 5, 2003 this sequence version replaced gi:30368690.

Contact: Genoscope National de Sequencage

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7766.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CSODK012AC120P1&amp;c=7766.r.

## FEATURES

source

Location/Qualifiers

1..959

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSODK012YB23"

/cell\_type="HELA CELLS COT 25-NORMALIZED"

/cell\_line="HELA"

/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match

Best Local Similarity 84.5%; Score 942; DB 5; Length 959;

Matches 953; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 6 GCCCGGGCGGCTCGACGAGAGGCGACCGAGATGGCGGAGCTTCAGCGG-CGGC 64

Db 6 GCCCGGGCGGCTCGACGAGAGGCGACCGAGATGGCGGAGCTTCAGCGGCGGC 65

Qy 65 GATCTCAGAGCTGCTCGCTGCTGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCC 124

Db 66 GATCTCAGAGCTGCTCGCTGCTGCTAAGCAAGATTGGAGTACTCGTCGTCACCTCC 125

Qy 125 AGTCGCGTAAGGTTGGCTGTCGCACTGCGGCCATTTGTGGATGGAACAGCGGGAGCAAG 184

Db 126 AGTCGCGTAAGGTTGGCTGTCGCACTGCGGCCATTTGTGGATGGAACAGCGGGAGCAAG 185

Qy 185 TGATCCCCCTGTGTGCGGGGATGGAACAGCTGCTCTCTAGAGATTGCTAACTGGAGAA 244

Db 186 TGATCCCCCTGTGTGCGGGGATGGAACAGCTGCTCTCTAGAGATTGCTAACTGGAGAA 245

Qy 245 CCACAGAGAGACTCTCAATACCAATACCAATGATGATGATGATGATGATGATGATGATGAT 304

Db 246 CCACAGAGAGACTCTCAATACCAATACCAATGATGATGATGATGATGATGATGATGATGAT 305

Qy 305 GGACATCTATGACAGTTGACGACCCATCTTAAGGCCATTTGCTGGAGGGGAGAGTGC 364

Db 306 GGACATCTATGACAGTTGACGACCCATCTTAAGGCCATTTGCTGGAGGGGAGAGTGC 365

Qy 365 CAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGCC 424

Db 366 CAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGCC 425

Qy 425 AGAGCAACTCGGGTGATCCCGCGGGCTCTCATGACCTCTCTGAGCTTCACAAGGAGGA 484

Db 426 AGAGCAACTCGGGTGATCCCGCGGGCTCTCATGACCTCTCTGAGCTTCACAAGGAGGA 485

Qy 485 GGGTCCGAGGCGCGGCGCATGGGCCCTTTCTGTGTCACCATGTCTTACCTAGAGATCTACCA 544







```
DEFINITION AL559769 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone CS0DG005YF11 5-PRIME, mRNA sequence.
ACCESSION AL559769
VERSION AL559769.3 GI:46185141
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31283900.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DG005YF11&c=7766.r.
FEATURES
Location/Qualifiers
1..999
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG005YF11"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 82.6%; Score 921.4; DB 1; Length 999;
Best Local Similarity 96.8%; Pred. No. 6.2e-243;
Matches 913; Conservative 22; Mismatches 8; Indels 0; Gaps 0;
QY 6 GCCCGGGCGGCTCAGCAGCAGAGGCGAGCGGAGTGGCGGAGCTTCAGCGCGCGG 65
DB 41 GCCCGGGCGGCTCAGCAGCAGAGGCGAGCGGAGTGGCGGAGCTTCAGCGCGCGG 100
QY 66 ATCTCAGGAGCTGGTGGCTGTCGGCTAAGCAAGATTGGAGTCTCGTCCACCTCCA 125
DB 101 ATCTCAGGAGCTGGTGGCTGTCGGCTAAGCAAGATTGGAGTCTCGTCCACCTCCA 160
QY 126 GCTCGGTAAGGGTGGCTGTCGAGTGGCGGCAATTTGGTGGTGAACAGCGGAGCAAGT 185
DB 161 GCTCGGTAAGGGTGGCTGTCGAGTGGCGGCAATTTGGTGGTGAACAGCGGAGCAAGT 220
QY 186 GATCCCCCTGTGTCGGGGGATGACAGTGTCTCTTAGAGATTGCTAACTGGAGGAAC 245
DB 221 GATCCCCCTGTGTCGGGGGATGACAGTGTCTCTTAGAGATTGCTAACTGGAGGAAC 280
QY 246 CACAGGAGACTCTCAATACCACTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 305
DB 281 CACAGGAGACTCTCAATACCACTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 340
QY 306 GACATCTATGAGGTTTCAAGTCAGCCCATCTTAAGGCATCTGCTGGAAGGCGAGATGCC 365
DB 341 GACATCTATGAGGTTTCAAGTCAGCCCATCTTAAGGCATCTGCTGGAAGGCGAGATGCC 400
QY 366 AGTGTGCTTGCTATGGACCCACAGGAGCTGGGAGAGCGCACCAATCTCGGCGAGCCCA 425
DB 401 AGTGTGCTTGCTATGGACCCACAGGAGCTGGGAGAGCGCACCAATCTCGGCGAGCCCA 460
```

```
QY 426 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGGAG 485
DB 461 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGGAG 520
QY 486 GGTGCGAGGGCGGCGCATGGGCCCTTTCTGTCAACATGCTTTACTAGAGATCTACCCAG 545
DB 521 GGTGCGAGGGCGGCGCATGGGCCCTTTCTGTCAACATGCTTTACTAGAGATCTACCCAG 580
QY 546 GAGAGGATTATAGACTCTCTGACCTGCTTGGGAGAGCTGGTATCCGAGAGAGTGC 605
DB 581 GAGAAGGATTATAGACTCTCTGACCTGCTTGGGAGAGCTGGTATCCGAGAGAGTGC 640
QY 606 CGGCGGAATATCTCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 665
DB 641 CGGCGGAATATCTCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 700
QY 666 GAGCGGCACTTCTCTCCAGCCAGTCAAAATCGGACTGTAGGAGCCACCCGGCTCAACAG 725
DB 701 GAGCGGCACTTCTCTCCAGCCAACTCGAAATCGGAGTGTAGGAGCCACCCGGCTCAACAG 760
QY 726 CGCTCCTCCGCGAGTCACTGCTGCTCTGCTCAAGGTGGACCGCGGGAACGTTTGGCC 785
DB 761 CGCTCCTCCGCGAATCAWRCCTGCTCTCTGCTCAAAGTGGACCGCGGGAACGTTWRRGC 820
QY 786 CCATTTGCGCAGGAGGAGGAACTCTACCTGATTGCTGCTGGGTGAGGAGCAAC 845
DB 821 CCATTTGCGCAGGAGGAGGAACTCTACCTGATTGCTGCTGGGTGAGGAGCAAC 880
QY 846 CGGCGCAGGCAACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
DB 881 CGGCGCAGGCAACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 940
QY 906 TTTGCTCTGGCAAGTGTAGTCCGCTGAATCAAGGCGCTCC 948
DB 941 TTTGCTCTGGCAAGTGTAGTCCGCTGAATCAAGGCGCTCC 983
RESULT 11
BX376199
LOCUS BX376199 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC024YA10 5-PRIME, mRNA sequence.
ACCESSION BX376199
VERSION BX376199.2 GI:46616791
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30452476..
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC024YA10&c=7766.r.
FEATURES
Location/Qualifiers
1..1094
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YA10"
FEATURES
source
```

```

/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

## ORIGIN

Query Match	82.6%	Score	921.4	DB	5	Length	1094
Best Local Similarity	99.2%	Pred.	No. 6.4e-243				
Matches	963	Conservative	4	Mismatches	0	Indels	4
Qy	6	GCCGCGGCGGCTCGACGACGAGCGACGAGATGGCGGAGCTTTCAGCGG-CGGC	64				
Db							
Qy	12	GCCGCGGCGGCTCGACGACGAGCGACGAGAGGCGGCGAGCTTTCAGCGGCGGC	71				
Db							
Qy	65	GATCTCAGGAGCTGGTGGCTGTCGGCTAAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCC	124				
Db							
Qy	72	GATCTCAGGAGCTGGTGGCTGTCGGCTAAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCC	131				
Db							
Qy	125	AGCTCGGCTAAGGGTGGCTGTGCACCTGCGGCCATTTGTGTGATGGAAACAGCGGGAGCAAG	184				
Db							
Qy	132	AGCTCGGCTAAGGGTGGCTGTGCACCTGCGGCCATTTGTGTGATGGAAACAGCGGGAGCAAG	191				
Db							
Qy	185	TGATCCCCCTGTGTGCGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAA	244				
Db							
Qy	192	TGATCCCCCTGTGTGCGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAA	251				
Db							
Qy	245	CCACGAGAGACTCTCAAAACCAAGTTTGATGCGCTTCTATGGGAGAGAGTACTCAGCA	304				
Db							
Qy	252	CCACGAGAGACTCTCAAAACCAAGTTTGATGCGCTTCTATGGGAGAGAGTACTCAGCA	311				
Db							
Qy	305	GGACATCTATGACGTTTCAGTGACGCCATCCTTAAGCACCTTGCTGGAAGGCGCAATGC	364				
Db							
Qy	312	GGACATCTATGACGTTTCAGTGACGCCATCCTTAAGCACCTTGCTGGAAGGCGCAATGC	371				
Db							
Qy	365	CAGTGTCTTGCCTATGACCCACAGAGAGCTGGGAAGCAGCACAAATGCTGGGCGAGCC	424				
Db							
Qy	372	CAGTGTCTTGCCTATGACCCACAGAGAGCTGGGAAGCAGCACAAATGCTGGGCGAGCC	431				
Db							
Qy	425	AGAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTGCACTCACAAGGGAGGA	484				
Db							
Qy	432	AGAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTGCACTCACAAGGGAGGA	491				
Db							
Qy	485	GGGTGCCGAGGCGCGCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACCA	544				
Db							
Qy	492	GGGTGCCGAGGCGCGCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACCA	551				
Db							
Qy	545	GGAGAGGTATTAGACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTG	604				
Db							
Qy	552	GGAGAGGTATTAGACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTG	611				
Db							
Qy	605	CCGGGGGAATATCTGTATTCGGGCTCTCTCCAGAAGCCCATCAGTAGCTTTTGCTGATTT	664				
Db							
Qy	612	CCGGGGGAATATCTGTATTCGGGCTCTCTCCAGAAGCCCATCAGTAGCTTTTGCTGATTT	671				
Db							
Qy	665	TGAGCGGCACTTCTTCGCAGCGAGTCGAAATTCGAGACTGTAGAGCCACCCGGCTCAACCA	724				
Db							
Qy	672	TGAGCGGCACTTCTTCGCAGCGAGTCGAAATTCGAGACTGTAGAGCCACCCGGCTCAACCA	731				
Db							
Qy	725	GCGCTCTCCGCGAGTATGCTGTGCTCCTGGTCAAGGTGACACAGCGGGAACTTTGGC	784				
Db							
Qy	732	GCGCTCTCCGCGAGTATGCTGTGCTCCTGGTCAAGGTGACACAGCGGGAACTTTGGC	791				
Db							
Qy	785	CCCATTTTCGCGAGCGAGGGGAAAACCTTACCTGATTGACTTGGCTGGGTGAGAGGACAA	844				
Db							
Qy	792	CCCATTTTCGCGAGCGAGGGGAAAACCTTACCTGATTGACTTGGCTGGGTGAGAGGACAA	851				
Db							
Qy	845	CCGCGCAGCAGGCAACAAAGGCGCTTCGGCT-AAAAGAGAGTGGAGCCATCAACACCTCCC	903				
Db							
Qy	852	CCGCGCAGCAGGCAACAAAGGCGCTTCGGCT-AAAAGAGAGTGGAGCCATCAACACCTCCC	911				
Db							
Qy	904	TGTTGTCTTCGGGCAAGTGGTAGATGGCTGGAATCAGGGGCTCCCTCGTGTACTTATC	963				
Db							

Db 912 TGTTTTGCTCGGCAAAAGTGGTAGATGGCTGAATCAGGG-CTCCCTCGTGTA-CTTATC 969

Qy 964 GGGACAGCAAG 974

Db 970 GGGACAGCAAS 980

RESULT 12

**LOCUS** BM910587 1031 bp mRNA linear EST 12-MAR-2002  
**DEFINITION** AGENCOURT\_6702754 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5454101  
5', mRNA sequence.  
**ACCESSION** BM910587  
**VERSION** BM910587.1 GI:19360966  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 1031)  
**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-x@mail.nih.gov](mailto:cgapbs-x@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

## FEATURES source

```

1. 1031
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5454101"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_98"
/notes="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II Rn (Life Technologies).
Note: this is a NIH MGC Library."

```

## ORIGIN

	Query Match	82.2%	Score 917;	DB 4;	Length 1031;
	Best Local Similarity	99.1%;	Pred. No. 1e-241;		
	Matches 931;	Conservative 0;	Mismatches 7;	Indels 1;	Gaps 1;
Qy	177	GGACCAAGTGATCCCCCTGTGTGCGGGGATGGACAGCTGCTCTCTAGAGATTGCTTAAC	236		
Db	1	GGACCAAGTGATCCCCCTGTGTGCGGGGATGGACAGCTGCTCTCTAGAGATTGCTTAAC	60		
Qy	237	TGGAGGAAACCAACGAGAGACTCTCAATAACAGTTTGATGCCCTTCTATGGGAGAGAGGT	296		
Db	61	TGGAGGAAACCAACGAGAGACTCTCAATAACAGTTTGATGCCCTTCTATGGGAGAGAGGT	120		
Qy	297	ACTCAGCAGGACATCTATGCAAGTTAGTCGACGCCCATCTTAGGCACTTGCTTGGAAAGG	356		
Db	121	ACTCAGCAGGACATCTATGCAAGTTAGTCGACGCCCATCTTAGGCACTTGCTTGGAAAGG	180		
Qy	357	CAGAAATGCCAGTGTGCTTGCTTATGGACCCACAGGAGCTGGGAAGACGCACACAATCTGT	416		
Db	181	CAGAAATGCCAGTGTGCTTGCTTATGGACCCACAGGAGCTGGGAAGACGCACACAATCTGT	240		

417 GCAGAGCCAGAGCAACTGGGGTGTATCCCGCGGGCTCTCATGAGACCTCTCTGAGCTCACA 476  
Db |  
241 GCAGAGCCAGAGCAACTGGGGTGTATCCCGCGGGCTCTCATGAGACCTCTCTGAGCTCACA 300  
Qy |  
477 AGGAGGAGGGTCCGAGGCGCGCCATGCGGCCCTTCTGTGACACATGCTTACCTAGAG 536  
Db |  
301 AGGAGGAGGGTCCGAGGCGCGCCATGCGGCCCTTCTGTGACACATGCTTACCTAGAG 360  
Qy |  
537 ATCTACAGAGAGGTATTAGACCTCTCGGACCTCTGCGGCCCTTCTGCGGAGACCTGCTTAATCCGA 596  
Db |  
361 ATCTACAGAGAGGTATTAGACCTCTCGGACCTCTGCGGCCCTTCTGCGGAGACCTGCTTAATCCGA 420  
Qy |  
597 GAAGACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTT 656  
Db |  
421 GAAGACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTT 480  
Qy |  
657 GCTGATTTGAGGGGACCTCTCGGACGAGTTCGAAATCGACTGTAGAGGCCACCGG 716  
Db |  
481 GCTGATTTGAGGGGACCTCTCGGACGAGTTCGAAATCGACTGTAGAGGCCACCGG 540  
Qy |  
717 CTCACACAGCGCTCTCCGCGAGTCTGCTGCTCTGCTCAAGGTGGACCGGGAA 776  
Db |  
541 CTCACACAGCGCTCTCCGCGAGTCTGCTGCTCTGCTCAAGGTGGACCGGGAA 600  
Qy |  
777 CGTTTGGCCCATTTTCGCGAGGAGGGAATACTTACTGATTTGACTTTGGCTGGGTCA 836  
Db |  
601 CGTTTGGCCCATTTTCGCGAGGAGGGAATACTTACTGATTTGACTTTGGCTGGGTCA 660  
Qy |  
837 GAGGACACCGCGGACAGGCAACAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAAC 896  
Db |  
661 GAGGACACCGCGGACAGGCAACAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAAC 720  
Qy |  
897 ACCTCCCTGTTGCTCGGCAAGTGTAGATGCGTGAATCAGGGCTCTCTGTTGA 956  
Db |  
721 ACCTCCCTGTTGCTCGGCAAGTGTAGATGCGTGAATCAGGGCTCTCTGTTGA 780  
Qy |  
957 CTTATCGGAGCAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCAC 1016  
Db |  
781 CTTATCGGAGCAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCAC 840  
Qy |  
1017 AGTATCTTTATGCCAATGTCCTCGAGAGACGCTTCTACCTAGACACAGTCTCCGCA 1076  
Db |  
841 AGTATCTTTATGCCAATGTCCTCGAGAGACGCTTCTACCTAGACACAGTCTCCGCA 900  
Qy |  
1077 CTCACCTTTGCTGCCAGG-TCCAAGGAGTGAATCAATTG 1114  
Db |  
901 CTCACCTTTGCTGCCAGGTTCCAAGGAGTGAATCAATCG 939

RESULT 13  
BM546877  
LOCUS BM546877 1019 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT\_6491037 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5723623  
5', mRNA sequence.  
ACCESSION BM546877  
VERSION BM546877.1 GI:18780186  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1019)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12711 row: j column: 08  
High quality sequence stop: 688.

## FEATURES

source

1..1019  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5723623"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_125"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site 1: EcoRV (destroyed); Site 2: NotI: RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

## ORIGIN

Query Match 81.5%; Score 909; DB 4; Length 1019;  
Best Local Similarity 97.5%; Pred. No. 1.7e-239;  
Matches 965; Conservative 0; Mismatches 21; Indels 4; Gaps 4;  
Qy 21 AGCAGCAGAGCGCAGCGAGATGCGCGAGCTTCAGCGCGCGGCGATCTCAGAGAGCTGGT 80  
Db 1 AGCAGCAGAGCGCAGCGAGATGCGCGAGCTTCAGCGCGCGGCGATCTCAGAGAGCTGGT 60  
Qy 81 CGCTCGCGCTAAGCAAGATTGGAGCTACTCGTCTGCCACCTCCAGCTCGCGCTAAGGGTG 140  
Db 61 CGCTCGCGCTAAGCAAGATTGGAGCTACTCGTCTGCCACCTCCAGCTCGCGCTAAGGGTG 120  
Qy 141 GCTGTGACCTCGCGCCATTTGTGGATGGAAAGCGGGAGCAAGTGAATCCCCCTCTGTGTG 200  
Db 121 GCTGTGACCTCGCGCCATTTGTGGATGGAAAGCGGGAGCAAGTGAATCCCCCTCTGTGTG 180  
Qy 201 CGGGCATGGAGCAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACAGAGAGACTCTC 260  
Db 181 CGGGCATGGAGCAGCTGCTCTTAGAGATTGCTAACTGGAGGAACACAGAGAGACTCTC 240  
Qy 261 AAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCTCAGCAGACATCTATCAGGT 320  
Db 241 AAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCTCAGCAGACATCTATCAGGT 300  
Qy 321 TCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGGGCAGAAATGCCAGTGTGTCTTCTAT 380  
Db 301 TCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGGGCAGAAATGCCAGTGTGTCTTCTAT 360  
Qy 381 GGACCCAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCAGAGCAACCTGGGGTG 440  
Db 361 GGACCCAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCAGAGCAACCTGGGGTG 420  
Qy 441 ATCCGCGGGCTCTCATGACCTCTCGAGCTCAAGGGAGGAGGCTGCCAGAGCGCGG 500  
Db 421 ATCCGCGGGCTCTCATGACCTCTCGAGCTCAAGGGAGGAGGCTGCCAGAGCGCGG 480  
Qy 501 CCATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACAGAGAGAGGTATTAGAC 560  
Db 481 CCATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACAGAGAGAGGTATTAGAC 540  
Qy 561 CTCTGGAGCCCTGCTTGGGAGACCTGTAATCCGAGAAGACTGCCGGGGGAATATCTCTG 620  
Db 541 CTCTGGAGCCCTGCTTGGGAGACCTGTAATCCGAGAAGACTGCCGGGGGAATATCTCTG 600  
Qy 621 ATTCCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGAGCGGACTCTCTG 680  
Db 601 ATTCCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGAGCGGACTCTCTG 660  
Qy 681 CCAGCCAGTGGAAATCGGACTGTAGGAGCCACCGCGCTCAACAGCGCTCTCTCCCGCAGT 740  
Db |

Db 661 CAGCCAGTGAATCGAGCTAGGAGCACCACCGGCTTACACAGCGCTCCTCCCGAGT 720  
 Qy 741 CATGCTGCTCTGCTGCTCAAGGTGGACCGAGGGAACGTTTGG-CGCCATTTCGCCAGCG 799  
 Db 721 CATGCTGCTCTGCTGCTCAAGGTGGACCGAGGGAACGTTTGGCCCCCAATTTCNGCAGCG 780  
 Qy 800 AGAGGGAACACTTACCTGATTGACTTGGCTGGGTGAGAGACACCGGCGCAGGCAA 859  
 Db 781 AGAGGGAACACTTACCTGATTGACTTGGCTGGGTGAGAGACACCGGCGCAGGCAA 840  
 Qy 860 CAAGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCCTGTTTGTCTCTGGGCAA 919  
 Db 841 CAAGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCCTGTTTGTCTGGGCAA 900  
 Qy 920 AGTGGTAGATCGCTGAATCAGGGCTCCTCTCGGTGA-CCTTATCGGGAAGCAA-GCTC 977  
 Db 901 AGTGGTAGATCGCTGAATCAGGGCTCCTCTCGGTGATACCCTAATCGGGAAGCAAAGGCTC 960  
 Qy 978 ACTGCGCTATTGCA-GGACTCTCTGGGTGG 1006  
 Db 961 ACTGCGCTATTGCAAGGAGCTCTCTGGGTGG 990

## RESULT 14

LOCUS BX400806 978 bp mRNA linear EST 28-APR-2004  
 DEFINITION BX400806 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
 cDNA clone CS0DK002YB19 5-PRIME, mRNA sequence.

ACCESSION BX400806

VERSION BX400806.2 GI:46845770

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE 1 (Bases 1 to 978)

JOURNAL Li W.B., Gruber C., Jesses, J. and Polayes, D.

COMMENT Full-length cDNA libraries and normalization

Unpublished (2001)

On May 13, 2003 this sequence version replaced gi:30614376.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

7766.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DK002CA10QPI&c=7766.r.

## FEATURES

Location/Qualifiers

1..978

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DK002YB19"

/cell\_type="HELA CELLS COT 25-NORMALIZED"

/cell\_line="HELA"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 81.4%; Score 907.8; DB 5; Length 978;  
 Best Local Similarity 99.2%; Pred No. 3, 5e-239;  
 Matches 951; Conservative 3; Mismatches 0; Indels 5; Gaps 4;  
 Qy 10 CGGCGGCTCGACGACGAGGCGACCGAGATCGGCGGCTTCACGCGGCGGATCT 69  
 |||||

Db 1 CGGGGGCTCGAGCGAGC--AGCGAGCGAGATGGCGGCA-STTCAGGGCGGCGATCT 57  
 Qy 70 CAGGAGCTGGTGGCTGGCTAAGCAAGATTGGAGCTACTGCTGCTCAGCTCCAGCTC 129  
 Db 58 CAGGAGCTGGTGGCTGGCTAAGCAAGATTGGAGCTACTGCTGCTCAGCTCCAGCTC 117  
 Qy 130 GCTTAAGGCTGGTGGCTGGCTAAGCAAGATTGGAGCTACTGCTGCTCAGCTCCAGCTC 189  
 Db 118 GCTTAAGGCTGGTGGCTGGCTAAGCAAGATTGGAGCTACTGCTGCTCAGCTCCAGCTC 177  
 Qy 190 CCCCTGTGTGGGGGCATGGACAGCTGCTCTAGAGATTGTAACTGGAGGAACACACC 249  
 Db 178 CCCCTGTGTGGGGGCATGGACAGCTGCTCTAGAGATTGTAACTGGAGGAACACACC 237  
 Qy 250 AGAGAGACTCTCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAAT 309  
 Db 238 AGAGAGACTCTCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAAT 297  
 Qy 310 TCTATGAGGTTTCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAAT 369  
 Db 298 TCTATGAGGTTTCAATACCAATACCAATACCAATACCAATACCAATACCAATACCAAT 357  
 Qy 370 TGTTCCTTATGACACCAAGAGCTGGAGAGCGACACATGCTTGGGAGGAGGAGTCTCAGCAGACA 429  
 Db 358 TGTTCCTTATGACACCAAGAGCTGGAGAGCGACACATGCTTGGGAGGAGGAGTCTCAGCAGACA 417  
 Qy 430 AACCTGGGGTGTATCCCGGGCTCTCATGAGACCTCTGAGCTCACAAGGAGGAGGAGTCT 489  
 Db 418 AACCTGGGGTGTATCCCGGGCTCTCATGAGACCTCTGAGCTCACAAGGAGGAGGAGTCT 477  
 Qy 490 CCGAGGGCGGGCCATGGGGCTTCTGTCTACCAATGCTTACCTAGAGATCTACCAAGAGA 549  
 Db 478 CCGAGGGCGGGCCATGGGGCTTCTGTCTACCAATGCTTACCTAGAGATCTACCAAGAGA 536  
 Qy 550 AGGTATTAGACCTCTCTGGAGACCTCTCTGGAGACCTCTGTAATCCGAGAGACTTCCCGGG 609  
 Db 537 AGGTATTAGACCTCTCTGGAGACCTCTCTGGAGACCTCTGTAATCCGAGAGACTTCCCGGG 596  
 Qy 610 GGAATATCTGATTCCCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTGTTTGGAGC 669  
 Db 597 GGAATATCTGATTCCCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTGTTTGGAGC 656  
 Qy 670 GGCACATCTCTCCAGCAGCTCGAAATCGGACTGTAGAGAGCCACCGGCTCAACAGCGCT 729  
 Db 657 GGCACATCTCTCCAGCAGCTCGAAATCGGACTGTAGAGAGCCACCGGCTCAACAGCGCT 716  
 Qy 730 CTTCGCCAGTCACTGCTGCTCTCTGTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 789  
 Db 717 CTTCGCCAGTCACTGCTGCTCTCTGTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 776  
 Qy 790 TTCCGCCAGCAGAGGGGAAACTCTTACCTGATTGACTTGGCTGGGTGAGAGGACAACCGGC 849  
 Db 777 TTCCGCCAGCAGAGGGGAAACTCTTACCTGATTGACTTGGCTGGGTGAGAGGACAACCGGC 836  
 Qy 850 GCACAGGCAACAAGGGCTTCCGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAG 909  
 Db 837 GCACAGGCAACAAGGGCTTCCGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAG 896  
 Qy 910 TCCTGGGC-AAAGTGTAGATGCGCTGAATCAGGGCTCCCTGCTGTACCTTATTCGGGA 967  
 Db 897 TCCTGGGC-AAAGTGTAGATGCGCTGAATCAGGGCTCCCTGCTGTACCTTATTCGGRM 955

## RESULT 15

LOCUS AL518361

DEFINITION AL518361 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone

CS0DA009YD01 5-PRIME, mRNA sequence.

ACCESSION AL518361

VERSION AL518361.3 GI:45654904

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

AL518361 1016 bp mRNA linear EST 23-MAR-2004

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1016)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 13, 2001 this sequence version replaced qi:305356600.  
COMMENT

	FEATURES	SOURCE
1.	100% Pure Cotton	Local Weavers
2.	Handloom Woven	Artisanal Craftsmen
3.	Eco-friendly Dyes	Natural Pigments
4.	Traditional Patterns	Cultural Heritage
5.	Durable & Long-lasting	Quality Materials
6.	Comfortable Fit	Expert Tailoring
7.	Versatile Styling	Modern Design
8.	Machine Washable	Easy Maintenance
9.	Available in Multiple Colors	Colorful Options
10.	Supports Local Economy	Community Empowerment

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA009YD01"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

Query Match	81.2%	Score 905;	DB 1;	Length 1016;
Best Local Similarity	98.5%;	Pred. No. 2.1e-238;		
Matches 931;	Conservative 3;	Mismatches 8;	Indels 3;	Gaps 2;
QY	70	CAGGAGCTGGTGGCTGGCTGGCTAAAGCAAGATTGGAGCTACTCGTGGTCCACCTCCAGCTC	129	
DB	1			
QY	130	GGCTAAGGGTGGCTGTGGCACTGCGGCCCATTTGTGGATGGAAACAGCGGGAGCAAGTGATC	189	
DB	61	GGCTAAGGGTGGCTGTGGCACTGCGGCCCATTTGTGGATGGAAACAGCGGGAGCAAGTGATC	120	
QY	190	CCCCCTGTGTGGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCACC	249	
DB	121	CCCCCTGTGTGGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCACC	180	
QY	250	AGGAGACTCTCAAAATACCAGTTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACA	309	
DB	181	AGGAGACTCTCAAAATACCAGTTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACA	240	
QY	310	TCTATGACAGTTTCAGTGCAGGCCCATCCTAAGGCACCTTGCTGGAAGGCGCAGAAATGCCAGTG	369	
DB	241	TCTATGACAGTTTCAGTGCAGGCCCATCCTAAGGCACCTTGCTGGAAGGCGCAGAAATGCCAGTG	300	
QY	370	TGCTTGCTTATGGAACCCACAGAGCTGGGAAGACGCACACAATGCTGGGCAGCCACAGAGC	429	
DB	301	TGCTTGCTTATGGAACCCACAGAGCTGGGAAGACGCACACAATGCTGGGCAGCCACAGAGC	360	
QY	430	AACTGGGGTGATCCCGGGGCTCTCATGGACTCTGTCAGCTCAACAGGGAGGAGGGTG	489	
DB	361	AACTGGGGTGATCCCGGGGCTCTCATGGACTCTGTCAGCTCAACAGGGAGGAGGGTG	420	
QY	490	CCGAGGCGCGGCATAGGGGCCCTTTCTGTGACCATGCTCTTACCTAGAGATCTACCAAGGAGA	549	
DB	421	CCGAGGCGCGGCATAGGGGCCCTTTCTGTGACCATGCTCTTACCTAGAGATCTACCAAGGAGA	480	
QY	550	AGGTATTAGACCTCTCGGACCTGCTTCGGGAGACCTGGTAAATCCGAGAAGACTGCCGGG	609	
DB	481	AGGTATTAGACCTCTCGGACCTGCTTCGGGAGACCTGGTAAATCCGAGAAGACTGCCGGG	540	

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 09:45:10 ; Search time 576.523 Seconds  
(without alignments)  
10152.424 Million cell updates/sec

Title: US-10-797-893-1  
Perfect score: 1115  
Sequence: 1 atgcagccgcggcggtctg.....ccaaggagtgatcaatga 1115

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1115	100.0	1115	10	Adc23337 DNA encod
2	1115	100.0	1115	12	Adg60227 Human mic
3	1112.4	99.8	1538	10	Adc23339 DNA encod
4	1112.4	99.8	1538	12	Adg60229 Human mic
5	1107.4	99.3	2097	12	Adg09241 Human KNS
6	1107.4	99.3	2099	10	Adk40995 Novel hum
7	1041	93.4	2104	3	Aaf15853 Human pro
8	1038	93.1	1041	10	Adc23341 DNA encod
9	1038	93.1	1041	12	Adg60231 Human mic
10	1035.4	92.9	1464	10	Adc23343 DNA encod
11	1035.4	92.9	1464	12	Adg60233 Human mic
12	265	23.8	464	9	Ach44493 Human toe
13	233.2	20.9	531	12	Ach77408 Human gen
14	230.4	20.7	232	12	Ach91108 Human gen
15	184.6	16.6	2095	11	Adm01564 Human cdn
16	184.6	16.6	3624	12	Ado44181 DNA encod
17	155.8	14.0	3374	10	Adj95073 Novel NOV
18	154.8	13.9	2375	8	Abx34596 Human mod
19	136.6	12.3	1152	6	Abx94616 Nucleotid
20	136.6	12.3	1152	8	Abx94616 DNA encod
21	136.6	12.3	4108	6	Abx94614 Human kin

22	136.6	12.3	4108	8	ABS57215	Abs57215 Human kin
23	135.2	12.1	2675	10	ADJ95077	Adj95077 Novel NOV
24	127.8	11.5	897	4	AAI70088	Aai70088 Human kin
25	127.8	11.5	897	6	AAL39619	Aal39619 DNA encod
26	127.8	11.5	897	12	ADM81051	Adm81051 Human Hsk
27	127.6	11.4	482	10	ABT40796	Abt40796 Toxicity
28	125.2	11.2	1026	6	ABQ73061	Abq73061 Human kin
29	125.2	11.2	1026	9	AAL56806	Aal56806 DNA encod
30	125.2	11.2	1026	10	ADB66785	Adb66785 DNA encod
31	124.2	11.1	1011	6	ABQ73062	Abq73062 Human kin
32	124.2	11.1	1011	9	AAL56807	Aal56807 DNA encod
33	124.2	11.1	1011	10	ADB66787	Adb66787 DNA encod
34	121.4	10.9	1014	6	ABA94615	Abx94615 Human Hsk
35	118.2	10.6	2132	10	ADC30338	Adc30338 Human nov
36	115	10.3	1839	10	ADC30569	Adc30569 Human nov
37	114.6	10.3	2034	4	ABL16733	Ab116733 Drosophil
38	114.6	10.3	3661	6	AD28568	Ad28568 Human kin
39	114.6	10.3	3694	10	ADC10189	Adc10189 Human NOV
40	114.6	10.3	4034	4	ABL16732	Ab116732 Drosophil
41	114	10.2	3570	12	ADQ19012	Adq19012 Human sof
42	113.6	10.2	987	6	AD28569	Ad28569 Human kin
43	110.8	9.9	4292	10	ADJ95085	Adj95085 Novel NOV
44	109.6	9.8	2394	4	ABL28591	Ab128591 Drosophil
45	104	9.3	1391	5	AAH89871	Aah89871 Human bon

ALIGNMENTS

RESULT 1  
ADC23337  
ID ADC23337 standard; DNA; 1115 BP.

AC ADC23337;

DT 18-DEC-2003 (first entry)

XX DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 1).  
XX human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.

OS Homo sapiens.

XX Key  
FH CDS  
FT Location/Qualifiers  
FT 1..1115  
FT /\*tag= a  
FT /product= "KID protein"  
FT /transl\_except= (pos: 1..5; aa: Met)  
FT /note= "This codon has an apparent 2 nucleotide insertion  
that alters the reading frame"

XX US6387644-B1.

PD 14-MAY-2002.

XX 28-NOV-2000; 2000US-00724224.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI: 2003-706919/67.

XX P-PSDB; ADC23338.

XX Identifying a candidate agent as modulator of function of a target  
XX protein for treating cellular proliferation disorders by adding a  
XX candidate agent to a mixture of the target protein that

PT	directly/indirectly produces ADP or phosphate.	
PS	Disclosure; SEQ ID NO 1; 26pp; English.	
XX		
CC	This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cytosolic, cardants, immunomodulators and antiinflammatory. Accordingly, through gene therapy, they can be used for the treatment of cancer, hyperplasia, restenosis, cardiac hypertrophy, immune disorders and inflammation. This polynucleotide sequence is human KID DNA (Seqid 1) encoding a full length KID enzyme of the invention.	
XX		
SQ	Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 1115; DB 10; Length 1115;	
	Best Local Similarity 100.0%; Pred. No. 1.6e-311;	
	Matches 1115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGCAGCGCGGGCGCTCGACGACGAGGACGCGGAGATGGCGGACGCTTCAGCGG 60	721 ACCAGCGCTCCTCCGCGAGTCATGCTGTCTCTCTGCTCAGGTCGACGCGGAAAGCTT 780
Db	1 ATGCAGCGCGGGCGCTCGACGACGAGGACGCGGAGATGGCGGACGCTTCAGCGG 60	721 ACCAGCGCTCCTCCGCGAGTCATGCTGTCTCTCTGCTCAGGTCGACGCGGAAAGCTT 780
Qy	61 CGGCGATCTCAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120	781 TGGCCCCATTTGGCCAGCGAGAGGAGGAAACTCTACCTGATTTGCTGGCTGGGTTCAGAGG 840
Db	61 CGGCGATCTCAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120	781 TGGCCCCATTTGGCCAGCGAGAGGAGGAAACTCTACCTGATTTGCTGGCTGGGTTCAGAGG 840
Qy	121 CTCAGCTCGCGTAAGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180	841 ACAACCGCGCACAGGCAACAAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCT 900
Db	121 CTCAGCTCGCGTAAGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180	841 ACAACCGCGCACAGGCAACAAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCT 900
Qy	181 CAAGTGATCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240	901 CCCTGTTTCTCTGGGCAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
Db	181 CAAGTGATCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240	901 CCCTGTTTCTCTGGGCAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
Qy	241 GGAACACACGAGGAGCTCTCAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTC 300	961 ATCGGAGACAGCAAGCTCACTCGCTATTTGAGGAGTCTCTGGGTGGTTCAGCCACAGTA 1020
Db	241 GGAACACACGAGGAGCTCTCAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTC 300	961 ATCGGAGACAGCAAGCTCACTCGCTATTTGAGGAGTCTCTGGGTGGTTCAGCCACAGTA 1020
Qy	301 AGCAGGACATCTATGCAAGTTTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360	1021 TCCTTATTCGCAATTTGGCCCTTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
Db	301 AGCAGGACATCTATGCAAGTTTCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360	1021 TCCTTATTCGCAATTTGGCCCTTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080
Qy	361 ATGCCAGTGTGCTTATGAGCCACGAGCTGGGAGCTGGGAGCTGGGAGCTGGGCA 420	1081 ACTTGTCTGCGAGTCCCAAGGAGTGATCAATTGA 1115
Db	361 ATGCCAGTGTGCTTATGAGCCACGAGCTGGGAGCTGGGAGCTGGGAGCTGGGCA 420	1081 ACTTGTCTGCGAGTCCCAAGGAGTGATCAATTGA 1115
Qy	421 GCCCAGAGCACTGGGGTGTATCCCGGGGCTCTCATGGACCTCTGAGCTCAACAGGG 480	RESULT 2
Db	421 GCCCAGAGCACTGGGGTGTATCCCGGGGCTCTCATGGACCTCTGAGCTCAACAGGG 480	ADQ60227
Qy	481 AGGAGGTGCGAGGCGCGGCTATGGGCGCTTCTGTACCATCTTACCTAGAGATCT 540	ID ADQ60227 standard; DNA; 1115 BP.
Db	481 AGGAGGTGCGAGGCGCGGCTATGGGCGCTTCTGTACCATCTTACCTAGAGATCT 540	XX ADQ60227;
Qy	541 ACCAGGAGAGGATATTAGCTTCTGAGCTGGGAGCTGGGAGCTGGTATCCGAGAG 600	AC AC
Db	541 ACCAGGAGAGGATATTAGCTTCTGAGCTGGGAGCTGGGAGCTGGTATCCGAGAG 600	DT 23-SEP-2004 (first entry)
Qy	601 ACTCCCGGGGAAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTACGTTTCTG 660	XX Human microtubule motor protein DNA #1.
Db	601 ACTCCCGGGGAAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTACGTTTCTG 660	XX Human; microtubule motor protein; gene; ds;
Qy	661 ATTTTGAAGCGCACTTCTGCGACGCGTGAATCGGAATCGGAGCCACCGGCTCA 720	KW cellular proliferation disorder; cancer; hyperplasia; restenosis;
Db	661 ATTTTGAAGCGCACTTCTGCGACGCGTGAATCGGAGCCACCGGCTCA 720	KW cardiac hypertrophy; immune disorder; inflammation;
		KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;
		KW graft rejection; inflammatory bowel disease; angioplasty.
		XX OS Homo sapiens.
		XX PN US6762043-B1.
		XX PD 13-JUL-2004.
		XX PF 06-MAR-2002; 2002US-00093317.
		XX PR 20-APR-1999; 99US-00295612.
		XX PR 20-JUN-2000; 2000US-00597292.
		XX PR 28-NOV-2000; 2000US-00724224.
		XX PA (CYTO-) CYTOKINETICS INC.
		XX PI Beraud C;
		XX PI WPI; 2004-532491/51.
		XX DR New isolated microtubule motor protein, useful for screening modulators
		XX PT for treating cellular proliferation disorders such as cancer,
		XX PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and
		XX PT inflammation.
		XX PS Disclosure; SEQ ID NO 1; 26pp; English.
		XX CC The invention relates to human microtubule motor proteins and the nucleic
		XX CC acids encoding them. The invention also relates to a method of screening
		XX CC for modulators of a motor protein which has microtubule stimulated ATPase



CC activity, a method of testing for ATPase activity of microtubule motor  
CC proteins, methods to identify candidate agents that bind to a target  
CC protein or act as a modulator of the binding characteristics or  
CC biological activity of a target protein, modulators of the target  
CC protein, and methods of treating cellular proliferation disorders such as  
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
CC and inflammation, for treating disorders associated with kinesin-like DNA  
CC binding protein (KID) and for inhibiting KID. The sequences are used for  
CC screening for modulators of motor proteins useful for treating cellular  
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
CC hypertrophy, immune disorders and inflammation, for treating disorders  
CC associated with KID and for inhibiting KID and for treating autoimmune  
CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents DNA encoding a human microtubule  
CC motor protein of the invention. Note: The specification states that this  
CC sequence encodes the protein featured as SEQ ID NO:2, but this does not  
CC appear to be the case.

SQ Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;

Query Match 100.0%; Score 1115; DB 12; Length 1115;  
Best Local Similarity 100.0%; Pred. No. 1.6e-311;  
Matches 1115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGCCGCGCGCTCGACGACGAGCGAGCGCGAGATGGCGGAGCTTCAGCGG 60  
DB 1 ATGCAGCCGCGCGCTCGACGACGAGCGAGCGCGAGATGGCGGAGCTTCAGCGG 60  
QY 61 CGCGGATCTCAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120  
DB 61 CGCGGATCTCAGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120  
QY 121 CTCACGCTCGGTAAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180  
DB 121 CTCACGCTCGGTAAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180  
QY 181 CAAAGTATCCCCCTGTGTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGA 240  
DB 181 CAAAGTATCCCCCTGTGTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGA 240  
QY 241 GGAACACACAGGAGCTCCTAAATACAGTTGATGCTTCTATGGGGAGAGTACTC 300  
DB 241 GGAACACACAGGAGCTCCTAAATACAGTTGATGCTTCTATGGGGAGAGTACTC 300  
QY 301 AGCAGACATCTATGAGCTTCAGTCAGCCCATCTTAAGGCACTTGTGGAGGGCACA 360  
DB 301 AGCAGACATCTATGAGCTTCAGTCAGCCCATCTTAAGGCACTTGTGGAGGGCACA 360  
QY 361 ATGCCAGTGTGCTTGCTATGACCCACAGGAGCTGGGAGAGCGCACAAATGCTGGCA 420  
DB 361 ATGCCAGTGTGCTTGCTATGACCCACAGGAGCTGGGAGAGCGCACAAATGCTGGCA 420  
QY 421 GCCCAGAGCAACTGGGGGTGATCCCGGGGTCTCATGACCTCTGACGCTCACAAGGG 480  
DB 421 GCCCAGAGCAACTGGGGGTGATCCCGGGGTCTCATGACCTCTGACGCTCACAAGGG 480  
QY 481 AGGAGGTGCGGAGGCGCGGCTTCTGTCACCATGCTTACCTAGAGATCT 540  
DB 481 AGGAGGTGCGGAGGCGCGGCTTCTGTCACCATGCTTACCTAGAGATCT 540  
QY 541 ACCAGAGAGGATATAGACCTCTCGGAGCTTGGGAGACCTGGTAATCCGAGAG 600  
DB 541 ACCAGAGAGGATATAGACCTCTCGGAGCTTGGGAGACCTGGTAATCCGAGAG 600  
QY 601 ACTGCCGGGGAATATCTGATTCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTG 660  
DB 601 ACTGCCGGGGAATATCTGATTCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTG 660  
QY 661 ATTTTGGGCGCACTTCTCGGAGCGGCTCGAATCGGACTGTAGGAGCCACCGGCTCA 720  
DB 661 ATTTTGGGCGCACTTCTCGGAGCGGCTCGAATCGGACTGTAGGAGCCACCGGCTCA 720

QY 721 ACCAGCGCTCTCCCGCAGTCATGCTGTGCTCCTGGTCAAGGTGGACCGGGAAGCTT 780  
DB 721 ACCAGCGCTCTCCCGCAGTCATGCTGTGCTCCTGGTCAAGGTGGACCGGGAAGCTT 780  
QY 781 TGGCCCCCATTTGCGCAGCAGAGGGAATACTTACCTGATTGACTTGGCTGGGTCAAGG 840  
DB 781 TGGCCCCCATTTGCGCAGCAGAGGGAATACTTACCTGATTGACTTGGCTGGGTCAAGG 840  
QY 841 ACAACCGCGCAGCAGCAACAAGGGCTTTCGGCTAAAGAGAGCTGGAGCCATCAACACCT 900  
DB 841 ACAACCGCGCAGCAGCAACAAGGGCTTTCGGCTAAAGAGAGCTGGAGCCATCAACACCT 900  
QY 901 CCTGTTTCTCTGGGCAAGTGGTAGATCGCTGAATCAGGGCCTCCCTCGTGTACCTT 960  
DB 901 CCTGTTTCTCTGGGCAAGTGGTAGATCGCTGAATCAGGGCCTCCCTCGTGTACCTT 960  
QY 961 ATCGGAGACGAGCTCACTCGCTTATTCAGAGCTCTCTGGTGGCTAGCCCAAGTA 1020  
DB 961 ATCGGAGACGAGCTCACTCGCTTATTCAGAGCTCTCTGGTGGCTAGCCCAAGTA 1020  
QY 1021 TCTTATTGCCAATTTGCCCTTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
DB 1021 TCTTATTGCCAATTTGCCCTTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
QY 1081 ACTTTGCTGCCAGGTCACAGGAGGTGATCAATTGA 1115  
DB 1081 ACTTTGCTGCCAGGTCACAGGAGGTGATCAATTGA 1115  
RESULT 3  
ADC23339  
ID ADC23339 standard; DNA; 1538 BP.  
XX  
AC ADC23339;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 3).  
XX  
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 1..1538  
FT /product= "KID protein"  
FT /transl\_except= (pos: 4..5; aa: Pro)  
FT /note= "This codon has an apparent 1 nucleotide deletion  
FT that alters the reading frame"  
XX  
PN US6387644-B1.  
XX  
XX 14-MAY-2002.  
XX  
XX 28-NOV-2000; 2000US-00724224.  
XX  
XX 20-APR-1999; 99US-00295612.  
XX 20-JUN-2000; 2000US-00597292.  
XX  
XX (CYTO-) CYTOKINETICS INC.  
XX  
XX Beraud C;  
XX  
XX WPI; 2003-706919/67.  
XX P-PSDB; ADC23340.  
XX  
XX Identifying a candidate agent as modulator of function of a target  
XX protein for treating cellular proliferation disorders by adding a  
XX candidate agent to a mixture of the target protein that



activity, a method of testing for ATPase activity of microtubule motor proteins, methods to identify candidate agents that bind to a target protein or act as a modulator of the binding characteristics or biological activity of a target protein, modulators of the target protein, and methods of treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with kinesin-like DNA binding protein (KID) and for inhibiting KID. The sequences are used for screening for modulators of motor proteins useful for treating cellular proliferation disorders such as cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation, for treating disorders associated with KID and for inhibiting KID and for treating autoimmune diseases, arthritis, graft rejection, inflammatory bowel disease and proliferation induced after medical procedures including surgery and angioplasty. This sequence represents DNA encoding a human microtubule motor protein of the invention. Note: The specification states that this sequence encodes the protein featured as SEQ ID NO:4, but this does not appear to be the case.

Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;

Query Match 99.8%; Score 1112.4; DB 12; Length 1538;

Query Match 99.8%; Score 1112.4; DB 12; Length 1538;

Best Local Similarity 99.9%; Pred. No. 1.1e-310;  
Matches 1113; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy	1	ATGACGCCCGCGGCTCGACGACAGAGCGGACGAGATGCGGAGCTTCAGCGG	60
Db	1	ATGACGCCCGCGGCGGCTCGACGACAGAGCGGACGAGATGCGGAGCTTCAGCGG	60
Qy	61	CGGCGATCTCAGGAGCTGGTCGTGCGGCTTAAGCAAGATTGGAGCTACTCGTCGTCCAC	120
Db	61	CGGCGATCTCAGGAGCTGGTCGTGCGGCTTAAGCAAGATTGGAGCTACTCGTCGTCCAC	120
Qy	121	CTCCAGGCTCGCGTAAGGGTGGCTGTGCGACTCGGCGCCATTTGTGTGATGCGAAACAGCGGGAG	180
Db	121	CTCCAGGCTCGCGTAAGGGTGGCTGTGCGACTCGGCGCCATTTGTGTGATGCGAAACAGCGGGAG	180
Qy	181	CAAGTGATGCCCTGTGTGCGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTCGA	240
Db	181	CAAGTGATGCCCTGTGTGCGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTCGA	240
Qy	241	GGAAACCACGAGAGACTCTCAAAATACAGTTTGTATGCTTCTATGGGAGAGGATCTC	300
Db	241	GGAAACCACGAGAGACTCTCAAAATACAGTTTGTATGCTTCTATGGGAGAGGATCTC	300
Qy	301	AGCAGGACATCTATGCAAGTTTCAGTCAGCGCCATCCTTAAGGCACTTGCTGGAAGGGCAGA	360
Db	301	AGCAGGACATCTATGCAAGTTTCAGTCAGCGCCATCCTTAAGGCACTTGCTGGAAGGGCAGA	360
Qy	361	ATGCCAGTGTGCTTGGCTATATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCA	420
Db	361	ATGCCAGTGTGCTTGGCTATATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCA	420
Qy	421	GCCCAGAGCAACCTGTGGGTGATCCCGGGGCTCTCATGGAACCTCTGCAGCTCAACAGG	480
Db	421	GCCCAGAGCAACCTGTGGGTGATCCCGGGGCTCTCATGGAACCTCTGCAGCTCAACAGG	480
Qy	481	AGGAGGTCGCGAGGGCGGCGCATGGGCGCTTCTGTGCACCATGCTTACCTAGAGATCT	540
Db	481	AGGAGGTCGCGAGGGCGGCGCATGGGCGCTTCTGTGCACCATGCTTACCTAGAGATCT	540
Qy	541	ACCAGGAGAAAGGTATTAGACCTCTCTGGACCTGTCTTCGGGAGACCTGTGTAATCCGAGAG	600
Db	541	ACCAGGAGAAAGGTATTAGACCTCTCTGGACCTGTCTTCGGGAGACCTGTGTAATCCGAGAG	600
Qy	601	ACTGCCGGGGAAATATCTGTATCCGGTCTCTCCCAAGAGCCCATCAGTAGCTTGTCTG	660
Db	601	ACTGCCGGGGAAATATCTGTATCCGGTCTCTCCCAAGAGCCCATCAGTAGCTTGTCTG	660
Qy	661	ATTTTGTAGCGGCATTCTCTGCAGCGCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCA	720
Db	661	ATTTTGTAGCGGCATTCTCTGCAGCGCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCA	720

Qy	721	ACGAGCGCTCTCCCGAGTCATGCTGTGCTCTCTGTCTCAAGGTGGACCAAGCGGGAACGTT	780
Db	721	ACCAAGCGCTCTCCCGCAGTCATGCTGTGCTCTGTGTCAAGGTGGACCAAGCGGGAACGTT	780
Qy	781	TGGCCCCATTTTGGCCACGAGAGGGAAAACTCTACCTGATTGACTTGGCTGGGTGCAGAGG	840
Db	781	TGGCCCCCATTTTGGCCACGAGAGGGAAAACTCTACCTGATTGACTTGGCTGGGTGCAGAGG	840
Qy	841	ACAAACGGCGCACAGGCACAAAGGGGCCCTTCGGGCTAAAAAGAGAGTGGAGCCATCAACACCT	900
Db	841	ACAAACGGCGCACAGGCACAAAGGGCCTTCGGGCTAAAGAGAGTGGAGCCATCAACACCT	900
Qy	901	CCCTGTTTGTCTGGGCAAAAGTGGTAGATCGGCTGAATCAAGGCGCTCCCTCGTGTACCTT	960
Db	901	CCCTGTTTGTCTGGGCAAAAGTGGTAGATCGGCTGAATCAAGGCGCTCCCTCGTGTACCTT	960
Qy	961	ATCGGACACAGCAAGCTACCTCGGCTATTGACAGAGCTCTCTGGGTGGCTCAGGCCACAGTA	1020
Db	961	ATCGGACACAGCAAGCTACCTCGGCTATTGACAGAGCTCTCTGGGTGGCTCAGGCCACAGTA	1020
Qy	1021	TCCTTATTGGCAACATTGGCCCTTGAGAGAGCGCTTCTACCTTAGACACAGTCTCCGCACTCA	1080
Db	1021	TCCTTATTGGCAACATTGGCCCTTGAGAGAGCGCTTCTACCTTAGACACAGTCTCCGCACTCA	1080
Qy	1081	ACTTTGTGCGCAGGTCCAAAGGAGGTGATCAATTG	1114
Db	1081	ACTTTGTGCGCAGGTCCAAAGGAGGTGATCAATCG	1114

RESULT 5  
ADO09241

ADQ09241  
ID ADQ09241 standard; cDNA; 2097 BP.

AA ADQ09241;

DT 23-SEP-2004 (first entry)

Human KNSL4 encoding cDNA SEQ ID NO:426.

KW thannatos-associated protein; THAP; THAP responsive gene; THAP family;  
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer  
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine  
 KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;  
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;  
 KW human; gene; ss.

AA  
OS  
Homo sapiens.

XX PN WO2004055050-A2.

XX  
PD 01-JUL-2004.XX  
PF 10-DEC-2003: 2003WO-IB006434.

XX  
PR 10-DEC-2002; 2002US-0432699P.

PR 03-JUL-2003; 2003US-0485027P.

AA  
PA  
(ENDO-) ENDOCUBE SAS.

PA (CNRS) CNRS CENT NAT RECH SCI.

PI Girard J, Amalric F, Roussigne M, Clouaire T;

DR WPI; 2004-525034/50.

DR P-PSDB; ADQ09240.

Modulating expression of a Thanatos (death)-Associated Protein (THAP) responsive gene for preventing or treating e.g. cancer or inflammation, comprises modulating the interaction of a THAP polypeptide with a nucleic acid.

XX PS Example 47; SEQ ID NO 426; 612pp; English.

XX  
CC The present invention describes a method for modulating the expression of





gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
wound; infectious disease, ss.  
Homo sapiens.  
W0200055174-A1.  
21-SEP-2000.  
08-MAR-2000; 2000WO-US005988.  
12-MAR-1999; 99US-0124270P.  
(HUMA-) HUMAN GENOME SCI INC.  
(ROSE/) ROSEN C A.  
Rosen CA, Ruben SM;  
WPI: 2000-587513/55.  
P-PSDB; AAB56650.  
Prostate cancer associated gene sequences, referred to as prostate cancer  
antigens, useful for treatment, prevention, and diagnosis of disorders  
such as prostate cancer.  
Claim 1; Page 805-806; 2338pp; English.  
AAF15566 to AAF16505 encode the human prostate cancer associated  
proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
The prostate cancer antigens can have neuroprotective, cytostatic,  
cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
nephrotropic, antiinfective, gynaecological and antibacterial activities,  
and can be used in gene therapy. The prostate cancer antigen  
polynucleotides may be used for detection of prostate cancer, chromosome  
identification, as chromosome markers, and for numerous other diagnostic  
or research purposes. The prostate cancer antigens may be used to treat  
disorders such as neural, immune, muscular, reproductive,  
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
AAF57303 represent sequences used in the exemplification of the present  
invention  
Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;  
Query Match 93.4%; Score 1041; DB 3; Length 2104;  
Best Local Similarity 99.7%; Pred. No. 5.2e-290;  
Matches 1041; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
71 AGGAGCTGGTGGCTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCG 130  
60 AGGAGCTGGTGGCTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCG 119  
131 CGTAAGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 190  
120 CGTAAGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 179  
191 CCCCTGTGGGGGCGATGGACAGTCTCTAGAGATTGCTAAGTGGAGGACACCA 250  
180 CCCCTGTGGGGGCGATGGACAGTCTCTAGAGATTGCTAAGTGGAGGACACCA 239  
251 GGAGACTCTCAATACAGATTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACAT 310  
240 GGAGACTCTCAATACAGATTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACAT 299  
311 CTATGCAAGTTTCAAGTGGACCCATCTTAAAGCACTTGTGGAAGGCGAAGTCCAGTGT 370  
300 CTATGCAAGTTTCAAGTGGACCCATCTTAAAGCACTTGTGGAAGGCGAAGTCCAGTGT 359  
371 GCTTGGCTATGACCCAGAGAGTGGGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 430  
360 GCTTGGCTATGACCCAGAGAGTGGGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 419  
431 ACCTGGGGTGATCCCGGGGCTCTCATGTGACCTCTCTGAGCTCAAGGGAGGGGTGC 490

Db 420 ACCTGGGGTGATCCCGGGGCTCTCATGACCTCTCTGAGTCACAAAGGAGGAGGTGC 479  
Qy 491 CGAGGGCGGGCCATGGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTTACAGAGAA 550  
Db 480 CGAGGGCGGGCCATGGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTTACAGAGAA 539  
Qy 551 GGTATTAGAGACTCTCTGAGCCCTGCTTCCGGAGACCTGGTAATCCGAGAGAGACTGCGCGGG 610  
Db 540 GGTATTAGAGACTCTCTGAGCCCTGCTTCCGGAGACCTGGTAATCCGAGAGAGACTGCGCGGG 599  
Qy 611 GAATATCTCTGATTCCGGGCTCTTCCAGAGGCCATCAGTGTGCTTGTGATTTGAGCG 670  
Db 600 GAATATCTCTGATTCCGGGCTCTTCCAGAGGCCATCAGTGTGCTTGTGATTTGAGCG 659  
Qy 671 GCACTTCTCTGAGCCAGTGGAAATCGGACTGTAGAGCCACCGGCTCAACAGCGCTC 730  
Db 660 GCACTTCTCTGAGCCAGTGGAAATCGGACTGTAGAGCCACCGGCTCAACAGCGCTC 719  
Qy 731 CTCCCGCAGTCATGCTGTGCTCTGCTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCATT 790  
Db 720 CTCCCGCAGTCATGCTGTGCTCTGCTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCATT 779  
Qy 791 TCGCCAGCGAGAGCGGAAACTCTTACCTGATTGACTTGGCTGGCTCAGAGCAACCGCGC 850  
Db 780 TCGCCAGCGAGAGCGGAAACTCTTACCTGATTGACTTGGCTGGCTCAGAGCAACCGCGC 839  
Qy 851 CACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGT 910  
Db 840 CACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGT 899  
Qy 911 CTTGGGCAAAAGTGTAGATGCGCTGAATCAGAGGCTCCCTCGTGTACTTATCGGACAG 970  
Db 900 CTTGGGCAAAAGTGTAGATGCGCTGAATCAGAGGCTCCCTCGTGTACTTATCGGACAG 959  
Qy 971 CAAAGTCACTCGCTATTGTCAGAGTCTCTGCGTGGCTAGCCACAGTATCTTATTGC 1030  
Db 960 CAAAGTCACTCGCTATTGTCAGAGTCTCTGCGTGGCTAGCCACAGTATCTTATTGC 1019  
Qy 1031 CAACATGCGCCCTGAGAGAGCTTCTACCTAGACAGTCTCCGACTCAACTTTGCTGC 1090  
Db 1020 CAACATGCGCCCTGAGAGAGCTTCTACCTAGACAGTCTCCGACTCAACTTTGCTGC 1079  
Qy 1091 CAGTCCAGGAGGTGATCAATTG 1114  
Db 1080 CAGTCCAGGAGGTGATCAATCG 1103  
RESULT 8  
ADC23341  
ID ADC23341 standard; DNA; 1041 BP.  
XX  
AC ADC23341;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE DNA encoding the human kinesin-like DNA binding protein (KID) (seqID 5).  
XX  
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 1..1041  
FT /\*tag= a  
FT /product= "KID protein"  
XX  
FN US6387644-B1.  
XX  
PD 14-MAY-2002.

XX 28-NOV-2000; 2000US-00724224.  
 PF PI  
 PR 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 XX (CYTO-) CYTOKINETICS INC.  
 PA  
 XX Beraud C;  
 PI  
 XX WPI: 2003-706919/67.  
 DR P-PSDB; ADC23342.  
 DR  
 XX Identifying a candidate agent as modulator of function of a target  
 PT protein for treating cellular proliferation disorders by adding a  
 PT candidate agent to a mixture of the target protein that  
 PT directly/indirectly produces ADP or phosphate.  
 XX  
 XX Disclosure; SEQ ID NO 5; 26pp; English.  
 XX  
 CC This invention relates to a novel method for high throughput screening  
 CC systems used to identify compounds for the treatment of cellular  
 CC proliferation disorders. Specifically, it refers to candidate agents that  
 CC are capable of modulating the activity of target proteins having motor  
 CC domains, such that the target protein directly or indirectly produces ADP  
 CC or phosphate. Furthermore, this activity can be determined using  
 CC fluorescence or absorbance readouts. The present invention describes a  
 CC method that identifies modulators of the target protein, which is a  
 CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardians,  
 CC immunomodulators and antiinflammatory. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasia,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polynucleotide sequence is human KID DNA (SeqID 5) encoding a full length  
 CC KID enzyme of the invention.  
 XX  
 SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;  
 Query Match 93.1%; Score 1038; DB 10; Length 1041;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-289;  
 Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GTCGCTGTCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAAG 137  
 DB 4 GTCGCTGTCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAAG 63  
 QY 138 GTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 197  
 DB 64 GTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCCTGT 123  
 QY 198 GTGGCGGGCATGACAGCTGTCTCTAGAGATTGCTTAAGTGGAGAACCAACAGGAGACT 257  
 DB 124 GTGGCGGGCATGACAGCTGTCTCTAGAGATTGCTTAAGTGGAGAACCAACAGGAGACT 183  
 QY 258 CTCAAATACAGTTTGTATGCTTCTATGGGGAGAGTACTCAGCAGACATCTATGCA 317  
 DB 184 CTCAAATACAGTTTGTATGCTTCTATGGGGAGAGTACTCAGCAGACATCTATGCA 243  
 QY 318 GGTTCAGTGCAGCCCATCTTAGGCACTTGTGGAGGCGCAATGCGAGTGTCTTCC 377  
 DB 244 GGTTCAGTGCAGCCCATCTTAGGCACTTGTGGAGGCGCAATGCGAGTGTCTTCC 303  
 QY 378 TATGACCCACAGGAGCTGGGAAGACGCACACATGCTGGGAGCCACAGAGCAACTGGG 437  
 DB 304 TATGACCCACAGGAGCTGGGAAGACGCACACATGCTGGGAGCCACAGAGCAACTGGG 363  
 QY 438 GTGATCCCGCGGCTCTCATGACCTCTCTGAGCTCACAAGGGAGAGGTGCCGAGGCG 497  
 DB 364 GTGATCCCGCGGCTCTCATGACCTCTCTGAGCTCACAAGGGAGAGGTGCCGAGGCG 423  
 QY 498 CGGCAATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 557  
 DB 424 CGGCAATGGGCGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 483

QY 558 GACCTCTGAGCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAAATATC 617  
 DB 484 GACCTCTGAGCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAAATATC 543  
 QY 618 CTGATTCCGGGCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACTTC 677  
 DB 544 CTGATTCCGGGCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGAAGCGGCACTTC 603  
 QY 678 CTGCGCAGCCAGTCGAAATCGGACTGTAGGAGCACCACCGGCTCAACAGCGCTCTCCCGC 737  
 DB 604 CTGCGCAGCCAGTCGAAATCGGACTGTAGGAGCACCACCGGCTCAACAGCGCTCTCCCGC 663  
 QY 738 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAACGTTTGGCCCCCATTTCCGCCAG 797  
 DB 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGAACGTTTGGCCCCCATTTCCGCCAG 723  
 QY 798 CGAGAGGGGAAACTCTACCTGATTGACTTTGGCTGGGTCAAGAGACAAACCGGCGCACAGGC 857  
 DB 724 CGAGAGGGGAAACTCTACCTGATTGACTTTGGCTGGGTCAAGAGACAAACCGGCGCACAGGC 783  
 QY 858 AACAAAGGGCCTTCGGCTAAACAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 917  
 DB 784 AACAAAGGGCCTTCGGCTAAACAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 843  
 QY 918 AAGGTGGTAGATCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 977  
 DB 844 AAGGTGGTAGATCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903  
 QY 978 ACTCGCTTATTCAGAGCTCTCTGGTGGCTCAGGCCACAGATGATCTTATGCGCAACTT 1037  
 DB 904 ACTCGCTTATTCAGAGCTCTCTGGTGGCTCAGGCCACAGATGATCTTATGCGCAACTT 963  
 QY 1038 GCCCTGTGAGAGCGCTTCTACCTAGACAGCTCTCGGCACTCAACTTTGCTGCCAGTCC 1097  
 DB 964 GCCCTGTGAGAGCGCTTCTACCTAGACAGCTCTCGGCACTCAACTTTGCTGCCAGTCC 1023  
 QY 1098 AAGGAGGTGATCAATTGA 1115  
 DB 1024 AAGGAGGTGATCAATTGA 1041

RESULT 9  
 ADQ60231  
 ID ADQ60231 standard; DNA; 1041 BP.  
 XX AC ADQ60231;  
 XX DT 23-SEP-2004 (first entry)  
 XX Human microtubule motor protein DNA #3.  
 DE Human; microtubule motor protein; gene; ds;  
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
 KW cardiac hypertrophy; immune disorder; inflammation;  
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;  
 KW graft rejection; inflammatory bowel disease; angioplasty.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 PH CDS 1..1041  
 FT /\*tag= a  
 FT /product= "Human microtubule motor protein #3"  
 XX  
 PN US6762043-B1.  
 XX 13-JUL-2004.  
 XX 06-MAR-2002; 2002US-00093317.  
 XX 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 PR 28-NOV-2000; 2000US-00724224.







XX Beraud C;  
 XX WPI: 2003-706919/67.  
 XX P-PSDB; AD23344.  
 PT Identifying a candidate agent as modulator of function of a target  
 PT protein for treating cellular proliferation disorders by adding a  
 PT candidate agent to a mixture of the target protein that  
 PT directly/indirectly produces ADP or phosphate.  
 XX Disclosure; SEQ ID NO 7; 26pp; English.  
 XX This invention relates to a novel method for high throughput screening  
 CC systems used to identify compounds for the treatment of cellular  
 CC proliferation disorders. Specifically, it refers to candidate agents that  
 CC are capable of modulating the activity of target proteins having motor  
 CC domains, such that the target protein directly or indirectly produces ADP  
 CC or phosphate. Furthermore, this activity can be determined using  
 CC fluorescence or absorbance readouts. The present invention describes a  
 CC method that identifies modulators of the target protein, which is a  
 CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardians,  
 CC immunomodulators and antiinflammatory. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polynucleotide sequence is human KID DNA (SeqID 7) encoding a full length  
 CC KID enzyme of the invention.  
 XX Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;  
 Query Match 92.9%; Score 1035.4; DB 10; Length 1464;  
 Best Local Similarity 99.9%; Pred. No. 1.9e-288;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 GGTCTGTCGCTAAGCAGATGGAGTACTGTCGTCACCTCCAGCTCGCTAAGG 137  
 DB 4 GGTCTGTCGCTAAGCAGATGGAGTACTGTCGTCACCTCCAGCTCGCTAAGG 63  
 QY 138 GTGGCTGTGCGACTGCGGCATTTGTGGATGAAACAGCGGAGCAAGTGTATCCCTGT 197  
 DB 64 GTGGCTGTGCGACTGCGGCATTTGTGGATGAAACAGCGGAGCAAGTGTATCCCTGT 123  
 QY 198 GTGCGGGGATGAGCAGCTGTCTCTAGAGATTGTCTAACTGGAGAAACACAGGAGACT 257  
 DB 124 GTGCGGGGATGAGCAGCTGTCTCTAGAGATTGTCTAACTGGAGAAACACAGGAGACT 183  
 QY 258 CTCATATACCACTTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317  
 DB 184 CTCATATACCACTTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243  
 QY 318 GGTTCAGTGCAGCCCATCTAAGGCATTTGTGGAGGAGGAGCAATGCCAGTGTGCTGCC 377  
 DB 244 GGTTCAGTGCAGCCCATCTAAGGCATTTGTGGAGGAGGAGCAATGCCAGTGTGCTGCC 303  
 QY 378 TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGCAACTGGG 437  
 DB 304 TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGCAACTGGG 363  
 QY 438 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAGGTCGCGAGGC 497  
 DB 364 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAGGTCGCGAGGC 423  
 QY 498 CGGCCATGGGCTTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 557  
 DB 424 CGGCCATGGGCTTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA 483  
 QY 558 GACCTCTTGGACCTCTGCTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGAAATATC 617  
 DB 484 GACCTCTTGGACCTCTGCTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGAAATATC 543  
 QY 618 CTGATTCGGGCTCTCTCCAGAGGCCATCAGTAGCTTTGTGATTTTGGCGGCACTTC 677  
 DB 544 CTGATTCGGGCTCTCTCCAGAGGCCATCAGTAGCTTTGTGATTTTGGCGGCACTTC 603

QY 678 CTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACCCAGCGCTCTCCCGC 737  
 DB 604 CTGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACCCAGCGCTCTCCCGC 663  
 QY 738 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACAGCGGGAAACGTTTGGCCCCCATTTGCCAG 797  
 DB 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACAGCGGGAAACGTTTGGCCCCCATTTGCCAG 723  
 QY 798 CGAGAGGGGAAACTCTACTGATTTGACTTGGCTGGGTGAGAGACACCGGGGCACAGGC 857  
 DB 724 CGAGAGGGGAAACTCTACTGATTTGACTTGGCTGGGTGAGAGACACCGGGGCACAGGC 783  
 QY 858 AACAGAGGCTCTCGGCTAAAGAGAGTGGAGCATCAACACCTCTCCCTGTTTGTCTGGGC 917  
 DB 784 AACAGAGGCTCTCGGCTAAAGAGAGTGGAGCATCAACACCTCTCCCTGTTTGTCTGGGC 843  
 QY 918 AAGTGGTAGATGCGCTGAATCAGGGCTCTCTGCTGTACCTTATCGGAGACAGCAAGTC 977  
 DB 844 AAGTGGTAGATGCGCTGAATCAGGGCTCTCTGCTGTACCTTATCGGAGACAGCAAGTC 903  
 QY 978 ACTCGCTTATTCAGGACTCTCTGCTGGTGGCTCAGCCACAGTATCTTATTCGCAACT 1037  
 DB 904 ACTCGCTTATTCAGGACTCTCTGCTGGTGGCTCAGCCACAGTATCTTATTCGCAACT 963  
 QY 1038 GCCCTGTAGAGAGCTTCTTACCTAGACAGCTCTCGGCACTCAACTTTGCTCGAGGTC 1097  
 DB 964 GCCCTGTAGAGAGCTTCTTACCTAGACAGCTCTCGGCACTCAACTTTGCTCGAGGTC 1023  
 QY 1098 AAGGAGGTGATCAATTG 1114  
 DB 1024 AAGGAGGTGATCAATCG 1040

RESULT 11  
 ADQ60233  
 ID ADQ60233 standard; DNA; 1464 BP.  
 XX AC ADQ60233;  
 XX DT 23-SEP-2004 (first entry)  
 XX DE Human microtubule motor protein DNA #4.  
 XX KW Human; microtubule motor protein; gene; ds;  
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
 KW cardiac hypertrophy; immune disorder; inflammation;  
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;  
 KW graft rejection; inflammatory bowel disease; angioplasty.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 CDS 1..1464  
 FT /\*tag= a  
 FT /product= "Human microtubule motor protein #4"  
 XX PN US6762043-B1.  
 XX PD 13-JUL-2004.  
 XX PF 06-MAR-2002; 2002US-00093317.  
 XX PR 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 PR 28-NOV-2000; 2000US-00724224.  
 XX PA (CYTO-) CYTOKINETICS INC.  
 XX PI Beraud C;  
 XX WPI: 2004-532491/51.  
 DR P-PSDB; ADQ60234.



XX The invention relates to an isolated polynucleotide comprising any one of  
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
 CC determined by the technique of SBH (sequencing by hybridisation). Also  
 CC included is a purified polypeptide comprising a sequence corresponding to  
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
 CC are useful in diagnostics as expressed sequence tags (EST) for  
 CC identifying expressed genes or for physical mapping of the human genome,  
 CC in forensics, in assessing biodiversity, or in identifying mutations  
 CC responsible for genetic disorders and other traits. The nucleotide  
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
 CC for chromosome and gene mapping, in the recombinant production of  
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
 CC is useful for generating antibodies specific for it. The present sequence  
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
 XX  
 SQ Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;

Query Match 23.8%; Score 265; DB 9; Length 464;  
 Best Local Similarity 77.6%; Pred. No. 4.9e-66;  
 Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;  
 QY 200 GCGGGCATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGAACCCACGAGAGACTCT 259  
 Db 37 GCTCGAATTCATATCGCGTGCATAGAGTGAACGCCGACCAACCCGATATCT 96  
 QY 260 CMAATACAGATTGATGCTTCTATGGGGAGAGAGTACTAGCAGGACATCTATGCGAGG 319  
 Db 97 CCCTTACCTGTGCTTGGCTTTGACAGCCATAGAGAGGACTGCGCCTGACATGTACTTCC 156  
 QY 320 TTCAGTCAGCCCATCTTAAGGCACTTCTGGAAGGCGAGATGCCAGTGTCTTGCCCTA 379  
 Db 157 TTGACCGTGGACCGTGCCTGCTGCTTCTTCTTAAGCGAGGAGCTGCGCCTA 216  
 QY 380 TGGACCCACAGAGCTGGGAGACGACACAAATGCTGGGCGAGCCGACGAGCAACCTGGGGT 439  
 Db 217 CGGTGCTGAGGAGCTGGCATTAC-ATCTGATGCTTGGCAGCCGACGAGCAACCTGGGGT 275  
 QY 440 GATCCCGGGGCTCTCATGAGACCTCTGAGCTCAACAGGAGGAGGTGCGGAGGCGG 499  
 Db 276 GATCCCGGGGCTCTCATGAGACCTCTGAGCTCAACAGGAGGAGGTGCGGAGGCGG 335  
 QY 500 GCCATGGGCGCTTCTGTACCATGCTTACCTAGATCTACCGAGGAGGATATTAGA 559  
 Db 336 GCCATGGGCGCTTCTGTACCATGCTTACCTAGATCTACCGAGGAGGATATTAGA 395  
 QY 560 CCTCTGGACCTGCTTGGGAGACCTGCTAATCCGAGAGACTGCCGGGGGAATATCT 619  
 Db 396 CCTCTGGACCTGCTTGGGAGACCTTGGATCCGAGAGACTGCCGGGGGAATATCT 455  
 QY 620 GATTCGGGG 628  
 Db 456 GATTCGGGG 464

RESULT 13  
 ACH77408  
 ID ACH77408 standard; DNA; 531 BP.  
 XX ACH77408;  
 AC ACH77408;  
 XX  
 XX 29-JUL-2004 (first entry)  
 DT Human genome derived single exon probe #10603.  
 XX Human; probe; ss; gene expression; single exon probe; microarray;  
 XX alternative splicing event; genomic alteration.  
 KW Homo sapiens.  
 XX  
 OS

PN US2003194704-A1.  
 XX 16-OCT-2003.  
 PD 03-APR-2002; 2002US-00029386.  
 XX 03-APR-2002; 2002US-00029386.  
 PF (PENN/) PENN S G.  
 XX (RANK/) RANK D R.  
 PA (HANK/) HANZEL D K.  
 PA  
 PI -Penn SG, Rank DR, Hanzel DK;  
 XX WPI; 2004-119264/12.  
 DR  
 XX New human genome-derived single exon nucleic acid probes useful for human  
 PT gene expression analysis, for identifying or characterizing alternative  
 PT splicing events, for assessing genomic alterations or as tools for  
 PT surveying tissues.  
 XX  
 XX Claim 15; SEQ ID NO 10603; 80pp; English.  
 PS The invention relates to a nucleic acid probe for measuring human gene  
 XX expression, comprising any of the 27,400 fully defined nucleotide  
 CC sequences in the specification, or their complements or fragments, and  
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
 CC fully defined in the specification. The probe is a single exon probe that  
 CC hybridises under high stringency conditions to a nucleic acid molecule  
 CC expressed in human cells or tissues. Also included are a spatially-  
 CC addressable set of single exon nucleic acid probes for measuring human  
 CC gene expression (comprising a plurality of single exon nucleic acid  
 CC probes cited above, where each of the plurality of probes is separately  
 CC and addressably isolatable or amplifiable from the plurality), a single  
 CC exon microarray for measuring human gene expression, a method of  
 CC measuring human gene expression, a vector comprising the single exon  
 CC probe cited above, an ORF-encoded peptide comprising at least 8  
 CC contiguous amino acids of any of the above-mentioned amino acid  
 CC sequences (optionally with conservative amino acid substitutions), an  
 CC isolated antibody that binds specifically to a peptide cited above,  
 CC methods of selling and/or licensing single exon probes or microarrays to  
 CC a customer desiring to measure gene expression, a method of providing  
 CC human gene expression data by subscription, and a computer-readable  
 CC storage medium which contains a database having a plurality of records  
 CC (each record including data on the expression of a single exon probe  
 CC cited above). The probe, methods and apparatus are useful in gene  
 CC expression analysis. The probes may be used as tools for surveying  
 CC tissues to detect the presence of expressed messages that contain their  
 CC specific exon, or in constructing genome-derived single exon microarrays.  
 CC In addition, the probes are used in identifying and characterising  
 CC alterations in the genomic locus that includes their exon, in assessing  
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
 CC or in expressing the ORF-encoded peptide. The present sequence is a human  
 CC single exon probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
 XX  
 SQ Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;

Query Match 20.9%; Score 233.2; DB 12; Length 531;  
 Best Local Similarity 98.7%; Pred. No. 8e-57;  
 Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 758 CAAGGTGGACGACGCGGGAACGTTTGGCCCCATTTGCGCAGCGAGAGGAACTTACCT 817  
 Db 34 CCAGGTGGACGACGCGGGAACGTTTGGCCCCATTTGCGCAGCGAGAGGAACTTACCT 93  
 QY 818 GATTGACTTGGCTGGGTCAGAGGACACCGCGCGCAGGCAACAGGCGCTTCGGCTAAA 877  
 Db 94 GATTGACTTGGCTGGGTCAGAGGACACCGCGCGCAGGCAACAGGCGCTTCGGCTAAA 153

QY 878 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGGTAGATGGCGTGAA 937  
DB 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGGTAGATGGCGTGAA 213  
QY 938 TCAGGGGCTCCCTCGTGTACCTTATCGGGACACGACGACGCTACGCGCTATTGCGAGAC 995  
DB 214 TCAGGGGCTCCCTCGTGTACCTTATCGGGACACGACGACGCTACGCGCTATTGCGAGTC 271

RESULT 14  
ACH91108  
ID ACH91108 standard; DNA; 232 BP.  
XX AC ACH91108;  
XX 29-JUL-2004 (first entry)  
XX Human genome derived single exon probe #24303.  
XX Human; probe; ss; gene expression; single exon probe; microarray;  
XX alternative splicing event; genomic alteration.  
XX Homo sapiens.  
XX US2003194704-A1.  
XX 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
XX (RANK/) RANK D R.  
XX (HANZ/) HANZEL D K.  
XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX New human genome-derived single exon nucleic acid probes useful for human  
XX gene expression analysis, for identifying or characterizing alternative  
XX splicing events, for assessing genomic alterations or as tools for  
XX surveying tissues.  
XX Claim 1; SEQ ID NO 24303; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 688 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridizes under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above). The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays.

CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX  
XX Sequence 232 BP; 55 A; 64 C; 68 G; 45 T; 0 U; 0 Other;

Query Match 20.7%; Score 230.4; DB 12; Length-232;  
Best Local Similarity 99.6%; Pred. No. 3.6e-56;  
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 762 GTGGACACGCGGGAACGTTTGGCCCCATTTCGCCAGCGAGGGGAAAACTTACCTGATT 821  
DB 1 GTGGACACGCGGGAACGTTTGGCCCCATTTCGCCAGCGAGGGGAAAACTTACCTGATT 60  
QY 822 GACTTGGCTGGTTCAGAGGACACCGCGGCACAGCAACAGGCGCTTCGGCTAAAGAG 881  
DB 61 GACTTGGCTGGTTCAGAGGACACCGCGGCACAGCAACAGGCGCTTCGGCTAAAGAG 120  
QY 882 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGGTAGATGGCGTGATCAG 941  
DB 121 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGGTAGATGGCGTGATCAG 180  
QY 942 GGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCTCGCTATTGCGAGG 993  
DB 181 GGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCTCGCTATTGCGAGG 232

RESULT 15  
ADM01564  
ID ADM01564 standard; cDNA; 2095 BP.  
XX AC ADM01564;  
XX 20-MAY-2004 (first entry)  
XX Human cDNA of the invention SEQ ID NO:249.  
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.  
XX Homo sapiens.  
XX EP1347046-A1.  
XX 24-SEP-2003.  
XX 12-APR-2002; 2002EP-00008400.  
XX 22-MAR-2002; 2002JP-00137785.  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX WPI; 2003-723558/69.  
XX P-PSDB; ADM04007.  
XX New polynucleotides and polypeptides are useful in gene therapy, for  
XX developing a diagnostic marker or medicines for regulating their  
XX expression and activity, or as a target of gene therapy.  
XX Claim 1; SEQ ID NO 249; 305pp; English.  
XX The invention relates to a novel human polynucleotide and the encoded  
XX polypeptide. A polynucleotide of the invention may have a use in gene  
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful

CC as a primer for synthesizing the polynucleotide or as a probe for  
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
 CC useful in gene therapy, for developing a diagnostic marker or medicines  
 CC for regulating their expression and activity, or as a target of gene  
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
 CC are useful as pharmaceutical agents. The present sequence represents a  
 CC cDNA sequence of the invention.

XX  
 SQ Sequence 2095 BP; 489 A; 603 C; 658 G; 345 T; 0 U; 0 Other;

Query Match 16.6%; Score 184.6; DB 11; Length 2095;  
 Best Local Similarity 53.7%; Pred. No. 1.6e-42;  
 Matches 463; Conservative 0; Mismatches 379; Indels 21; Gaps 3;

QY	264	TACCAAGTTCGAGTCTTCTATGGGAGAGAGATCTCAGCAGGACATCTATGACGTTCA	323
DB	286	TACCTGTTTCGACGTGGCCTTTGACTTCACCGCCACCCAGGAGATGGTGTATCAGGCCACC	345
QY	324	GTGACGCCCATCTTAAGGCATTTGCTGGAAGGCGAGATGCCAGTGTGCTTGCCTATGGA	383
DB	346	ACCAAGAGCCTCATCGAGGGCTCATCTCAGGCTCAATGCCACTGTCTTTGCCCTATGSC	405
QY	384	CCACAGAGCTGGAGAGCGACACATGCTGGGCGCCAGAGCAACCTGGGGTGATC	443
DB	406	CCACAGGCTGGGAAACCTACACCATGCTGGGCGCAGACCAGGAGCCTGGCATCTAT	465
QY	444	CGCGGGCTCTCATGACCTCTCGAGCTCACAGGGAGGGTGGCGAGGGCGGCCA	503
DB	466	GTTCAGACCTCAAGACCTCTTCGTGCCATCGAGGAGACCAGCAATGA	516
QY	504	TGGGCCCCCTTTGTCACCATGCTTTACCTAGAGATCTACAGGAGAAGGTATTAGACCTC	563
DB	517	ATGGAGTATGAGGTCTCCATGCTCTACCTGGAGATCTCAATGAGATGATCCGGGACCTG	576
QY	564	CTGGACCTCTGCTCGGAGACCTGTGTAATCCAGAGACTGCCGGGGGAATATCTGATT	623
DB	577	TGAAACCCCTCCCTGGGCTACCTGGAGCTGGGGAGGACTCTAAGGGGGTGATCCAGGTG	636
QY	624	CGGGTCTCTCCAGAACCCCATAGTAGCTTTGCTGATTTTGGCGGCACCTCTCTGCCA	683
DB	637	GCCGCAATCACCGAAGTCTCCACCATCAATGCCAAGGAGATCATGAGCTGCTGATGAAG	696
QY	684	GCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGCAGTCA	743
DB	697	GGGAACCGCAGAGGACCCAGGAGCCACCGCGCCCAACCCAGACGCTCTCCGCTCCAC	756
QY	744	GCTGTCTCTGCTCAAGGTGACACAGCGGGAACGTTTGGCCCCAATTTGCCA-----G	797
DB	757	GCGGTACTGCAAGGTGACCGTGGCCGACGCGCGGTCAGAACATCTTGCAGGAGGTG	816
QY	798	CGAGAGGGAAMAACTCTACTGATTGACTTGGCTGGGTGAGAGCAACCGGCGCACAGGC	857
DB	817	CGGAGGGCGGCTGTTTATGATGCACTGGCTGGCTCAGAGCGGCTTCGAGACACAG	876
QY	858	AACAAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGGC	917
DB	877	AATCGTGGCGAGCGTATGAAGGAGGGGCCACATCAACCGCTCACTGCTGSCACTGGGC	936
QY	918	AAAGTGGTAGTGCCTG-----AATCAGGGCCTCCCTCGTGTACCTTATCGGGACAGC	971
DB	937	AACATCATCAAGCGCTGAGCGACAAAGGTTAGCAACAAGTATCACTCAACTATCGCGACGC	996
QY	972	AAGCTCACTCGCTATTTCAGGACTCTCTGGTGGCTCAGGCCACAGTATCCTTATTGCC	1031
DB	997	AHGCTACCCCGGCTCCTGAAGGACTCTCTGGAGAGAAACAGCCGCACAGTATGATCGCT	1056
QY	1032	AACATTGCCCTTGAGAGCGCTTCTTACCTTAGACACAGTCTCGGACTCAACTTTTGTGCC	1091
DB	1057	CACATCAGTCTCTGAGGAGTGCCTTCGAGGAGTCCCGGAACACCTGACCTACGCCGCG	1116
QY	1092	AGGTCGAAGGAGGTGATCAATTG	1114
DB	1117	CGGGCCAAAGAACTTAAGACTAG	1139

Search completed: November 10, 2004, 06:57:50  
 Job time : 582.523 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:25:02 ; Search time 5089.37 Seconds  
(without alignments)  
10482.189 Million cell updates/sec

Title: US-10-797-893-7  
Perfect score: 1464  
Sequence: 1 atgggtcgtctgcgctaag.....accattgtcccaaatgtga 1464

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1454.8	99.4	2016	3	CR590954 full-leng
2	1454.8	99.4	2025	3	CR590974 full-leng
3	1454.8	99.4	2057	3	CR615781 full-leng
4	1067.8	72.9	2055	3	AK084310
5	1063.6	72.7	2056	3	AK075922 Mus muscu
6	1007.6	68.8	1124	1	AL545712 AL545712
7	973.8	66.5	1027	5	EX337406 BX337406
8	970.6	66.3	1032	5	EX337353 BX337353
9	922.4	63.0	1031	4	BM910587 AGENCOURT
10	912.8	62.3	1083	1	AL559772
11	897.6	61.3	1016	1	AL518361
12	882.8	60.3	1034	5	BQ051730 AGENCOURT
13	881.6	60.2	959	5	BX363358 BX363358
14	875.4	59.8	971	5	BQ053671 AGENCOURT
15	874.5	59.7	958	5	EX354166 BX354166
16	866.2	59.2	978	5	EX400806 BX400806
17	860.8	58.8	1094	5	EX376199 BX376199
18	854.4	58.4	926	1	AL578338
19	852.5	58.2	1019	4	BM546877 AGENCOURT
20	851.4	58.2	865	6	CD109092 AGENCOURT
21	849.4	58.0	999	1	AL559769
22	841.6	57.5	915	5	BX383181 BX383181
23	823.6	56.3	1139	4	BM553107 AGENCOURT
24	821.4	56.1	1640	3	AK088033 Mus muscu

25	820.8	56.1	978	4	BI754589
26	819.4	56.0	981	4	BM561588 AGENCOURT
27	817.8	55.9	1051	4	BM806996 AGENCOURT
28	817.6	55.8	980	5	EX404364 BX404364
29	817.2	55.8	1041	4	BM478215 AGENCOURT
30	813	55.5	908	5	BQ953445 AGENCOURT
31	808.4	55.2	878	1	AL576578 AL576578
32	804.8	55.0	1040	4	BM546897 AGENCOURT
33	802.6	54.8	845	4	BQ686012 602638562
34	795.6	54.3	865	1	AL555550 AL555550
35	790.6	54.0	855	5	EX372162 BX372162
36	789.6	53.9	991	5	EX333336 BX333336
37	786	53.7	1103	4	BM458841 AGENCOURT
38	782.4	53.4	995	1	AL559752 AL559752
39	782.2	53.4	905	5	EX387403 BX387403
40	782	53.4	921	5	EU860011 AGENCOURT
41	773.8	52.9	913	5	EX445789 BX445789
42	772	52.7	990	5	EX399192 BX399192
43	768.6	52.5	846	5	EX404325 BX404325
44	767.4	52.4	1142	4	BM918115 AGENCOURT
45	761	52.0	838	5	EX341048 BX341048

#### ALIGNMENTS

RESULT 1  
CR590954  
LOCUS  
DEFINITION  
full-length cDNA clone CS0DI040YI06 of Placenta Cot 25-normalized  
of Homo sapiens (human).  
ACCESSION  
CR590954  
VERSION  
CR590954.1 GI:50471761  
KEYWORDS  
HTC; CNSLT\_CDNA.  
SOURCE  
Homo sapiens  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 2016)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
JOURNAL  
Contact: Feng Liang Email: fliang@lifetech.com URL:  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue  
Genoskope.  
2 (bases 1 to 2016)  
Direct Submission  
Submitted (20-JUL-2004) Genoskope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoskope.cns.fr  
- Web : www.genoskope.cns.fr)  
COMMENT  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
FEATURES  
Location/Qualifiers  
source  
1..2016  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI040YI06"  
/tissue\_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"  
ORIGIN  
Query Match 99.4%; Score 1454.8; DB 3; Length 2016;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GGTGCTGTCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 63  
|||||  
Db 9 GGTGCTGTCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 68  
|||||

QY	64	GTGGCTGTGCGACTCGGCCATTTTGTGGATGGAACAGCGGGAGCAAGTGAATCCCCCTGT	123	QY	1144	GAGCAAGAGGAGATTGGAGCCCTGAGCCATGCGAGCTCCAGGCTCTGCTCCAGAAA	1203
DB	69	GTGGCTGTGCGACTCGGCCATTTTGTGGATGGAACAGCGGGAGCAAGTGAATCCCCCTGT	128	DB	1149	GAGCAAGAGGAGATTGGAGCCCTGAGCCATGCGAGCTCCAGGCTCTGCTCCAGAAA	1208
QY	124	GTGCGGGGATGGAACAGCTGCTCTTAGAGATTGCTTAAGTGAAGAACCCAGGAGACT	183	QY	1204	CTCAGCCCTTACAGAAAGCTAAGCAGCATGACCCGCGCATGCTGGAGCGCTCTCTCAGC	1263
DB	129	GTGCGGGGATGGAACAGCTGCTCTTAGAGATTGCTTAAGTGAAGAACCCAGGAGACT	188	DB	1209	CTCAGCCCTTACAGAAAGCTAAGCAGCATGACCCGCGCATGCTGGAGCGCTCTCTCAGC	1268
QY	184	CTCAATACAGTTTATGCTTCTATGAGGAGAGGAGTACTCAGCAGGACATCTATGCA	243	QY	1264	TTGAGACCTGTGCTTGTCTCCAGGGAGCAGGGGGCCCTCTGTTGAGTACCCCAAG	1323
DB	189	CTCAATACAGTTTATGCTTCTATGAGGAGAGGAGTACTCAGCAGGACATCTATGCA	248	DB	1269	TTGAGACCTGTGCTTGTCTCCAGGGAGCAGGGGGCCCTCTGTTGAGTACCCCAAG	1328
QY	244	GGTTCACTGAGCCCAATCTTAAGGCATTTGCTGGAAGGCGCAAGATGCTGCTGCTG	303	QY	1324	CGAGCGGATGCTGCTTAATGAACACAGTGAAGAGAGGACCTAGAGATTGAGAGGCTT	1383
DB	249	GGTTCACTGAGCCCAATCTTAAGGCATTTGCTGGAAGGCGCAAGATGCTGCTGCTG	308	DB	1329	CGAGCGGATGCTGCTTAATGAACACAGTGAAGAGAGGACCTAGAGATTGAGAGGCTT	1388
QY	304	TATGGAACCAAGAGCTGGGAAGACGACACATGCTGGGAGCCAGCAACTGGG	363	QY	1384	AAGCAAGCAAGCAAGCAAGCAAGTGGAGCCAGAGATGTTGGCCAGAGGCTGAGGAAAGAG	1443
DB	309	TATGGAACCAAGAGCTGGGAAGACGACACATGCTGGGAGCCAGCAACTGGG	368	DB	1389	AAGCAAGCAAGCAAGCAAGCAAGTGGAGCCAGAGATGTTGGCCAGAGGCTGAGGAAAGAG	1448
QY	364	GTGATCCCGGGCTCTCATGAGACCTCTGACGCTCAAGGAGGAGGCTGCGAGGGC	423	QY	1444	AACCATTTGCCCAATG	1461
DB	369	GTGATCCCGGGCTCTCATGAGACCTCTGACGCTCAAGGAGGAGGCTGCGAGGGC	428	DB	1449	AACCATTTGCCCAATG	1466
QY	424	CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGAGAGGATTA	483	RESULT 2			
DB	429	CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGAGAGGATTA	488	CRS90974			
QY	484	GACCTCTGAGACCTGCTTGGGAGACCTGTTAATCGAGAGACTGCGGGGGAATATC	543	LOCUS			
DB	489	GACCTCTGAGACCTGCTTGGGAGACCTGTTAATCGAGAGACTGCGGGGGAATATC	548	DEFINITION			
QY	544	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC	603	ACCESSION			
DB	549	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC	608	VERSION			
QY	604	CTGCGAGCCAGTGAATCGACTGTAGAGGACACCGGCTCAACAGCGCTCTCCCGC	663	KEYWORDS			
DB	609	CTGCGAGCCAGTGAATCGACTGTAGAGGACACCGGCTCAACAGCGCTCTCTCCCGC	668	SOURCE			
QY	664	AGTCATGCTGTGCTTCTGCTCAAGTGGACACCGGGAAGCTTTGGCCCATTTGCCAG	723	ORGANISM			
DB	669	AGTCATGCTGTGCTTCTGCTCAAGTGGACACCGGGAAGCTTTGGCCCATTTGCCAG	728	REFERENCE			
QY	724	CGAGAGGGAACCTCTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGCGCGCAGGC	783	AUTHORS			
DB	729	CGAGAGGGAACCTCTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGCGCGCAGGC	788	TITLE			
QY	784	AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	843	JOURNAL			
DB	789	AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	848	REMARK			
QY	844	AAAGTGTGATGCTGATCAAGGAGTGGAGGCTCCCTGCTGCTTATCGGAGCAGCAAGCTC	903	COMMENT			
DB	849	AAAGTGTGATGCTGATCAAGGAGTGGAGGCTCCCTGCTGCTTATCGGAGCAGCAAGCTC	908	FEATURES			
QY	904	ACTCGCTTATGAGGACTCTCTCGGTGGGTGAGCCCAAGTATCTTATGCGCAACATT	963	source			
DB	909	ACTCGCTTATGAGGACTCTCTCGGTGGGTGAGCCCAAGTATCTTATGCGCAACATT	968	1..2025			
QY	964	GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC	1023	/organism="Homo sapiens"			
DB	969	GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC	1028	/mol_type="mRNA"			
QY	1024	AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCATGCTTGGGACCT	1083	/mol_xref="taxon:9606"			
DB	1029	AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCATGCTTGGGACCT	1088	/clone="CS01009Y001"			
QY	1084	GTTAAGCTGTCTAGAAAGAAATTTGCTTCCAGAGGCAAGAGAGCCCGAGGCTT	1143	/issue type="Placenta Cot 25-normalized"			
DB	1089	GTTAAGCTGTCTAGAAAGAAATTTGCTTCCAGAGGCAAGAGAGCCCGAGGCTT	1148	/plasmid="pWVSPORT_6"			

Query Match 99.4%; Score 1454.8; DB 3; Length 2025;  
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
 Matches 1456; Conservative

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCCGTAAGG 63  
 DB 59 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCCGTAAGG 118



QY 64 GTGGCTGTGCGAGCTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGT 123  
Db 119 GTGGCTGTGCGAGCTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGT 178  
QY 124 GTGGGGGCATGGAGCAGCTGCTCTAGAGATTGCTAACTGAGGAACCAACAGAGACT 183  
Db 179 GTGGGGGCATGGAGCAGCTGCTCTAGAGATTGCTAACTGAGGAACCAACAGAGACT 238  
QY 184 CTCAAATACCAATGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243  
Db 239 CTCAAATACCAATGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 298  
QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGCTGGAAGGCGAGAAATGCCAGTGTGCTGCC 303  
Db 299 GGTTCAGTGCAGCCCATCTTAAGGCACTTGCTGGAAGGCGAGAAATGCCAGTGTGCTGCC 358  
QY 304 TATGACCCACAGGAGCTGGAAGACGACACAAATGCTGSGCAGCCGACCCAGCAACCTGGG 363  
Db 359 TATGACCCACAGGAGCTGGAAGACGACACAAATGCTGSGCAGCCGACCCAGCAACCTGGG 418  
QY 364 GTGATCCCGCGGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGGAGGTGCCAGGGC 423  
Db 419 GTGATCCCGCGGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGGAGGTGCCAGGGC 478  
QY 424 CGGCATGGGCGCTTCTGTCTCAATGCTTTACCTAGAGATCTTACAGGAGAAAGTATTA 483  
Db 479 CGGCATGGGCGCTTCTGTCAACATGCTTTACCTAGAGATCTTACAGGAGAAAGTATTA 538  
QY 484 GACCTCTGGACCTGCTTGGGAGACTGCTTAATCCGAGAGACTGCGGGGGAATATC 543  
Db 539 GACCTCTGGACCTGCTTGGGAGACTGCTTAATCCGAGAGACTGCGGGGGAATATC 598  
QY 544 CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGTAGCGGCATTC 603  
Db 599 CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGTAGCGGCATTC 658  
QY 604 CTGCGACGAGTGAATGCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663  
Db 659 CTGCGACGAGTGAATGCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 718  
QY 664 AGTCATCTGTGCTGCTGCTCAAGTGTGAGCGGAGCCAGCGGGAAGCTTTGGGCCCATTTCCGCAG 723  
Db 719 AGTCATCTGTGCTGCTGCTCAAGTGTGAGCGGAGCCAGCGGGAAGCTTTGGGCCCATTTCCGCAG 778  
QY 724 CGAGAGGAAACCTCTACCTGATTGACTTTGGCTGCGGTGAGAGCAACCGCGCACAGGC 783  
Db 779 CGAGAGGAAACCTCTACCTGATTGACTTTGGCTGCGGTGAGAGCAACCGCGCACAGGC 838  
QY 784 AACAGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 843  
Db 839 AACAGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 898  
QY 844 AAGTGTAGATGGCTGAATCAGGCGCTCTGCTGTACTTATCGGAGCAGCAAGCTC 903  
Db 899 AAGTGTAGATGGCTGAATCAGGCGCTCTCTGCTGTACTTATCGGAGCAGCAAGCTC 958  
QY 904 ACTCGCTATTGAGGACTCTCTGGTGGCTCAGCCACAGTAGTATCTTATTTGCCAATT 963  
Db 959 ACTCGCTATTGAGGACTCTCTGGTGGCTCAGCCACAGTAGTATCTTATTTGCCAATT 1018  
QY 964 GCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1023  
Db 1019 GCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1078  
QY 1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGAGCTCTATGCTTGGGACCT 1083  
Db 1079 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGAGCTCTATGCTTGGGACCT 1138  
QY 1084 GTTAAGCTGTCTAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1143  
Db 1139 GTTAAGCTGTCTAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1198  
QY 1144 GAGGAAGAGAGATTTGGAGCCCTTGAGCCCATGCGAGCTCCAGCTCTGCTCTCCAGAAA 1203

Db 1199 GAGGAAGAGAGATCGGGAGCCCTGAGCCCTGAGCCCTCTGCGCTCTCCAGAAA 1258  
QY 1204 CTCAGCCCTTACAGAGCTAAGCAGACATGACCCGGGCATGCTGGAGGCGCTCTCAGC 1263  
Db 1259 CTCAGCCCTTACAGAGCTAAGCAGACATGACCCGGGCATGCTGGAGGCGCTCTCAGC 1318  
QY 1264 TTGACCGTCTGCTTGGCTCCAGGGAGCCAGGGGGCCCTCTGCTTGCAGTACCCCAAG 1323  
Db 1319 TTGACCGTCTGCTTGGCTCCAGGGAGCCAGGGGGCCCTCTGCTTGCAGTACCCCAAG 1378  
QY 1324 CGAGAGCGGATGCTGCTTAATGAAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1383  
Db 1379 CGAGAGCGGATGCTGCTTAATGAAGACAGTGAAGAGAGGACCTTAGAGATTGAGAGGCTT 1438  
QY 1384 AAGACGAGCAAAAAGAACTGGAGCCCAAGATGTTGGCCAGAGGCTGAGAGAAAGAG 1443  
Db 1439 AAGACGAGCAAAAAGAACTGGAGCCCAAGATGTTGGCCAGAGGCTGAGAGAAAGAG 1498  
QY 1444 AACATTGTCCCAATG 1461  
Db 1499 AACATTGTCCCAATG 1516

RESULT 3  
CR615781  
LOCUS 2057 bp mRNA linear HTC 21-JUL-2004  
DEFINITION full-length cDNA clone CSODK012YE23 of HeLa cells Cot 25-normalized  
of Homo sapiens (human).  
ACCESSION CR615781  
VERSION CR615781.1 GI:50496588  
KEYWORDS HTC; CNSLT cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2057)  
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paraday Avenue  
2 (bases 1 to 2057)  
Genoscope.  
Direct Submission

TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

FEATURES  
source Location/Qualifiers  
1..2057  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODK012YE23"  
/tissue\_type="HeLa cells Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

ORIGIN  
Query Match 99.4%; Score 1454.8; DB 3; Length 2057;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAGATTGGAGCTACTGCTGCTCCACTCCAGCTCGGTAAGG 63  
Db 79 GGTGCTGTGCGCTAAGCAGATTGGAGCTACTGCTGCTCCACTCCAGCTCGGTAAGG 138  
QY 64 GTGCTGTGCGACTGCGGCCATTTGTGGATGGAACAGCGGAGCAAGTGATCCCCCTGT 123

Db	139	GTGGCTGTGCGACTGGGCCATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGT	198	Db	1219	GAGGAAGAGAGATCGGGAGCCCTGAGCCCATGGCAGCTCCAGCTCTGCTCCCAAAA	1278
Qy	124	GTGCGGGGATGAGACGTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT	183	Qy	1204	CTCAGCCCCCTACAGAAAGCTTAAGCAGATGAGATCGACCCGGCCATGCTGGAGCGCTCTCAGC	1263
Db	199	GTGCGGGGATGAGACGTGCTCTCTAGAGATTGCTTAAGTGGAGAACACACAGGAGACT	258	Db	1279	CTCAGCCCCCTACAGAAAGCTTAAGCAGATGAGATCGACCCGGCCATGCTGGAGCGCTCTCAGC	1338
Qy	184	CTCAATACACAGTTTGTATGATGCTTCTATGGGAGAGAGTACTCTCAGCAGGACATCTATGCA	243	Qy	1264	TTGACCGCTGCTGCTCTCCAGGGGAGCAGGGGGCCCTCTGTTGAGTACCCCAAG	1323
Db	259	CTCAATACACAGTTTGTATGATGCTTCTATGGGAGAGAGTACTCTCAGCAGGACATCTATGCA	318	Db	1339	TTGACCGCTGCTGCTCTCCAGGGGAGCAGGGGGCCCTCTGTTGAGTACCCCAAG	1398
Qy	244	GGTTCACTAGTCAGCCCATCTAAGGCATTTGCTGGAAGGCGAGATGCTTGGTTCCTGCTG	303	Qy	1324	CGAGCGCGATGCTGCTTAATGAAGACAGTGAAGAGAAAGACCTAGAGATTGAGAGCTT	1383
Db	319	GGTTCACTAGTCAGCCCATCTAAGGCATTTGCTGGAAGGCGAGATGCTTGGTTCCTGCTG	378	Db	1399	CGAGCGCGATGCTGCTTAATGAAGACAGTGAAGAGAAAGACCTAGAGATTGAGAGCTT	1458
Qy	304	TATGGACCCACAGGAGCTGGAGAGCGACACAACTGCTGGGAGCCAGAGCAACTGGG	363	Qy	1384	AAGACGAGCAAAAGAACTGGAGGCCAAGATGTTGCCCGCAGAGGCTGAGGAAAAGAG	1443
Db	379	TATGGACCCACAGGAGCTGGAGAGCGACACAACTGCTGGGAGCCAGAGCAACTGGG	438	Db	1459	AAGACGAGCAAAAGAACTGGAGGCCAAGATGTTGCCCGCAGAGGCTGAGGAAAAGAG	1518
Qy	364	GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCCTGAGCTCACAAGGAGGAGGCTCCAGGGC	423	Qy	1444	AACCATTTGCTCCCAATG 1461	
Db	439	GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCCTGAGCTCACAAGGAGGAGGCTCCAGGGC	498	Db	1519	AACCATTTGCTCCCAATG 1536	
Qy	424	CGGCCATGGGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTACAGGAGAGTATTA	483				
Db	499	CGGCCATGGGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTACAGGAGAGTATTA	558	RESULT 4			
Qy	484	GACCTCTGAGCCCTGCTCGGAGAGCTGCTGTAATCCGAGAGACTGCGGGGGATATC	543	AK084310	2055 bp	mRNA	linear HTC 03-APR-2004
Db	559	GACCTCTGAGCCCTGCTCGGAGAGCTGCTGTAATCCGAGAGACTGCGGGGGATATC	618	LOCUS	Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length		
Qy	544	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTGTCTGATTTGAGCGGCACTTC	603	DEFINITION	enriched library, clone:D230020105 product:kinesin superfamily		
Db	619	CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTGTCTGATTTGAGCGGCACTTC	678	ACCESSION	protein KIF22, full insert sequence.		
Qy	604	CTGCACCCAGTCGAATTCGATGTAGGAGCCACCGGCTCAACAGAGCTCTCTCCCG	663	VERSION	AK084310.1 GI:26101874		
Db	679	CTGCACCCAGTCGAATTCGATGTAGGAGCCACCGGCTCAACAGAGCTCTCTCCCG	738	KEYWORDS	HTC; CAP trapper.		
Qy	664	AGTCATGCTGCTCTCTGCTCAAGTGGAGCCAGCGGAACTGTTGGCCCCATTTGCGCAG	723	SOURCE	Mus musculus (house mouse)		
Db	739	AGTCATGCTGCTCTCTGCTCAAGTGGAGCCAGCGGAACTGTTGGCCCCATTTGCGCAG	798	ORGANISM	Mus musculus		
Qy	724	CGAGAGGAAACTCTACCTGATTTGACTTGGCTGGGTTCAGAGACACACCGGCGACAGGC	783	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Db	799	CGAGAGGAAACTCTACCTGATTTGACTTGGCTGGGTTCAGAGACACACCGGCGACAGGC	858	1	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
Qy	784	AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	843	2	Carninci, P. and Hayashizaki, Y.		
Db	859	AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	918	AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
Qy	844	AAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACTTATCGGACAGCAAGCTC	903	TITLE	High-efficiency full-length cDNA cloning		
Db	919	AAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACTTATCGGACAGCAAGCTC	978	JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
Qy	904	ACTGCTTATTCAGGACTCTCGGTGGCTCAGCCACACATCTCTTATTGCCAACATT	963	MEDLINE	99279253		
Db	979	ACTGCTTATTCAGGACTCTCTGGTGGCTCAGCCACAGATCTCTTATTGCCAACATT	1038	PUBMED	10349636		
Qy	964	GCCCCCTAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTGTGCGCAGTCC	1023	REFERENCE	10349636		
Db	1039	GCCCCCTAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTGTGCGCAGTCC	1098	AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
Qy	1024	AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGCTCATGCTTGGGACCT	1083	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
Db	1099	AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGCTCATGCTTGGGACCT	1158	4	Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,		
Qy	1084	GTTAAGCTGCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT	1143	5	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
Db	1159	GTTAAGCTGCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT	1218	TITLE	Functional annotation of a full-length mouse cDNA collection		
Qy	1144	GAGGAAGGAGATTGGAGGCCCTGAGCCCATGGAGCTCAGGCTCTGCTCCCAAAA	1203	JOURNAL	Nature 409, 685-690 (2001)		
				AUTHORS	The RIKEN Consortium and the RIKEN Genome Exploration Research		
				REFERENCE	Group Phase I & II Team.		
					Analysis of the mouse transcriptome based on functional annotation		
					of 60,770 full-length cDNAs		

JOURNAL  
REFERENCE  
AUTHORS

Nature 420, 563-573 (2002)  
6 (bases 1 to 2055)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, O., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, I., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

TITLE  
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.jp/>  
URL: <http://fantom.gsc.riken.jp/>

## FEATURES

## source

Location/Qualifiers  
1..2055  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM,DB:D230020105"  
/db\_xref="taxon:10090"  
/clone="D230020105"  
/tissue\_type="eyeball"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="12 days embryo"  
8..1931  
misc\_feature  
/note="kinesin superfamily protein KIF22  
putative"

## ORIGIN

Query Match 72.9%; Score 1067.8; DB 3; Length 2055;  
Best Local Similarity 84.1%; Pred. No. 1.2e-273;  
Matches 1204; Conservative 0; Mismatches 227; Indels 0; Gaps 0;  
Qy 6 TCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCGTCCACCTCCAGCTCGCGTAAGGTT 65  
Db 70 TGGCGGCTGTGTAGCAAGAGAGGGCTTGGCGGGCTCGGCTCTAGCTGAGTAAGGTT 129  
Qy 66 GGCTGTGGCACTCGCGGCATTGTGTGATGGAAACAGCGGGAGCAAGTATCCCGCTGTGT 125  
Db 130 GGCTGTGCGGTGCGCGCGTTTATGTGATGGAGACAGACAGCGAAGGAGTCTCCCTGTGT 189  
Qy 126 GCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACACAGGAGACTCT 185  
Db 190 CCGAGCCATAGACAGCTGCTCTCTTGAAGTGGCTTAACCTGGAAAGAAATACACAGGAGAGCT 249  
Qy 186 CAAATACAGTTTATGTCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGG 245  
Db 250 CAAATATCAGTTTATGTCCTTCTATGGGAGAGGACACTCAGCAGGAGTCTATGATGG 309  
Qy 246 TTCAGTGCAGCCCATCTTAAGCACTTCTGTGAAAGGCGAGAAATGCCAGTGTGTGCTTA 305  
Db 310 TTCAGTACAGCCCATCTTAAGCACTTCTGTGAAAGGCGAGAAATGCCAGTGTGTGCTTA 369  
Qy 306 TGGACCCACAGGAGCTGGAGAGCGCACATGCTGGGAGCGCCACAGAGCACTGGGGT 365  
Db 370 TGGGCTTACTTGGGGCAGGAGAGACACACAATGCTGGGAGCGCCAGAGCAACCTGGAGT 429

Qy 366 GATCCCGCGGGCTCTCATGGACCTCCTCGAGCTCAAGGGAGGGGTGCCAGGGCG 425  
Db 430 GATTCTCTGGGCTCTCATGGACCTCCTCGAAGTACAGGAGGAGAGAGTGTGTGAGGAGC 489  
Qy 426 GCCATGGGCGCTTCTGTGTCACATGTCTTACCTAGAGATCTACAGGAGAAAGTATTAGA 485  
Db 490 GCCATGGGACGCTTCTGTGTGCTCTATTTAGAGATCTACACAGGAAAGGATTAGA 549  
Qy 486 CCTCTGGACCTGCTTCGGGAGACCTGGTAACTCGAGAGAGCTCCCGGGGAAATATCTCT 545  
Db 550 CCTCTGGATCCTGATCAGGAGACCTCGTATCCGGAAGACTCCCGAGGGAACATCTCT 609  
Qy 546 GATTCCGGGTCTCTCCCAAGAGCCCATCAGTAGCTTTCTGATTTTGTAGCGGCATCTCTCT 605  
Db 610 GATCCAGGCTCTACACAGAGCCCATCACCAGCTTCTCTGACTTCGAGCAGCATCTCTCT 669  
Qy 606 GCCAGCAGTGCAGAAATCGAAGTGTAGGAGCCACCGGCTCAACACAGCGCTCTCTCCGAG 665  
Db 670 TCCAGCCAGTAGAAATCGAGCTGTAGGAGCCACCGGCTTAAACAGCGCTCTCTCCGAG 729  
Qy 666 TCATGCTGCTCTCTGCTCAAGGTGGACCGGAGGAGCTTTGGCCCCATTTCCGCCAGG 725  
Db 730 TCACGAGTGTCTTGTGTCAGGTAGATCAGCGTGAACGTTTGAATCTCCATTTCCGCCAG 789  
Qy 726 AGAGGAGAAACTCTACCTGATTGACTTGGCTGGGTTCAGAGGACCAACCGCGGCACAGGCA 785  
Db 790 GGAAGGAAACTCTACCTTATGATTGGCTGGTTCAGAGGACCAACCTTCGACAGGCA 849  
Qy 786 CAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTCTTTGTCTGGGCA 845  
Db 850 CCAGGCAATTCGGCTCAAGAGAGCGGAGCCATCAACCTCCCTCTTTGTACTGGGCA 909  
Qy 846 AGTGTAGATGCGCTGATCAGGGCTCCCTCTGTTGTTTATCGGAGCAGCAAGCTCAC 905  
Db 910 AGTGTGATGTCATTAAACAGGGCTCCCTCTCGCATACCATACCGGAGCAGCAAGCTCAC 969  
Qy 906 TCGCTATTGTCAGGACTCTCTGGGTGCTCAGCCACAGTATCTTATTTGCCAATGTC 965  
Db 970 TCGTCTCTGTCAGAGACTCTCTGGGAGGCTCAGCTCATAGCATCTCTATGCCAATTTGC 1029  
Qy 966 CCCTGAGAGAGCTTTTACCTPAGACACAGTCTCCGCACTCAACTTTGTCTGCGAGTCCAA 1025  
Db 1030 TCCTGAAAGACGTTTTTACCAGGATACAAATTCAGCATTTGAATTCCTGCTAGTCCAA 1089  
Qy 1026 GGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCGCTCATGCTTTGGACCTGT 1085  
Db 1090 GGAGTGTATTAACCGGCTTTTCAACCAATGAGAGTCTACAGCTCATGCTTTGGCACCTGT 1149  
Qy 1086 TAAGCTGTCTCAGAAAGAAATTTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCGCTGA 1145  
Db 1150 TAAGCTGTCTCAGAAAGAAATTTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCGCTGA 1209  
Qy 1146 GGAAGAGGAGATTGGGAGCGCTGAGCCCATGCGAGCTTCCAGCTCTGCTCTCCAGAACT 1205  
Db 1210 AGAAGAAATCAACTGGGAGCGCTGAGTCTACAGAGCTCCAGCGTCTGCTCTCCAGAACT 1269  
Qy 1206 CAGCCCCCTACAGAGCTAAGCAGCATGGAGCCCGGCGCATGCTGGAGCGCTCTCTCAGGTT 1265  
Db 1270 AAGCCTCTTACAGAAAGCTAAGCAATATGGATCCAGCCATGTTGGAGAACTCTCTGAGTAT 1329  
Qy 1266 GGACCGTCTGCTTGGCTCCAGGGAGCGCCAGGGGCGCTCTGTTGTAGTACCCCAAGCG 1325  
Db 1330 GGAAGCTTGTCTGGGTTCCCAAGGAGCGCAAGGAGACCCCTACTGTAATATCCCAAGCG 1389  
Qy 1326 AGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTTAA 1385  
Db 1390 AGAAGAAATGGTGTCTATGAAGACAGTGGAGGAAAGAACTTATAGAGATTGAGAGGCTTAA 1449  
Qy 1386 GACGAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCCCAGAGGCTGAGGA 1436  
Db 1450 GNTGAAGCAAAAGAACTGGAAGCAAGGTGTTGGCCCCAGGAGCGACCGGA 1500



QY 726 AGAGGAACTCTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGGCGCACAGGCAA 785  
 Db |||||  
 QY 793 GGAAGAACTCTACATTAATTGATTTGGCTGGTTGAGAGCAACCGTGCACAGGCAA 852  
 Db |||||  
 QY 786 CAAGGGCTTGGCTAAAGAGAGTGGAGCATCAACACCTCCCTTTGTTGCTGGGCAA 845  
 Db |||||  
 QY 853 CCAGGGCATCGGCTCAAGAGAGCGGAGCCATCAACACCTCCCTCTTTGTTGCTGGGCA 912  
 Db |||||  
 QY 846 AGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTTATCGGAGCAGCAAGCTCAC 905  
 Db |||||  
 QY 913 AGTGTGATGATTAACACAGGGCTCCCTCGCATACCATACCGGAGACCAAGCTCAC 972  
 Db |||||  
 QY 906 TCGCTATTTGAGGACTCTCTGGGTGGCTGACGCCACAGTATCTTTATGCAACATTGC 965  
 Db |||||  
 QY 973 TCGTCTGCTGAGGACTCTCTGGGAGGCTCAGCTCATAGCATCCTCATTTGCCAATGTC 1032  
 Db |||||  
 QY 966 CCCTGAGAGAGCTTCTACCTAGACACAGTCTCGCACTCAACTTTGCTGTCAGGTCAA 1025  
 Db |||||  
 QY 1033 TCCTGAAAGAGCTTTTACCAAGGATACAAATTCAGCATTTCACTGCTAGGTCCAA 1092  
 Db |||||  
 QY 1026 GGAGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCCTCATGCTTTGGACCTGT 1085  
 Db |||||  
 QY 1093 GGAGTGATTAACCGGCTTTTACCAATGAGAGTCTACAGCCTCATGCTTTGGACCTGT 1152  
 Db |||||  
 QY 1086 TAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGAGGCAAGAGAGCCGAGCCCTGA 1145  
 Db |||||  
 QY 1153 TAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGAGGCAAGAGAGCCGAGCCCTGA 1212  
 Db |||||  
 QY 1146 GGAAGAGAGATTTGGAGCCCTGAGCCATGCGAGCTCCAGCCTCTGCTCCCAAGAACT 1205  
 Db |||||  
 QY 1213 AGAAGATCAACTGGGAGCCCTGAGTCTACAGAGCTCCAGCCTCTGCTCCCAAGAACT 1272  
 Db |||||  
 QY 1206 CAGCCCTCTACAGAGCTTAAGCAGCATGAGCCCGGCAATGCTGAGCGCCTCTCAGCTT 1265  
 Db |||||  
 QY 1273 AAGCTCTTACAGAGCTTAAGCAATNTGATCCAGCCATGTTGAGAACCTCTCTGAGTAT 1332  
 Db |||||  
 QY 1266 GGACCGTCTGCTTCCCTCCAGGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCAAGCG 1325  
 Db |||||  
 QY 1333 GGAACGTTTGTGGGTTCCAGGAGCAAGCAAGGAGCCCTCTACTGAATACCCCAAGCG 1392  
 Db |||||  
 QY 1326 AGACCGATGTTGCTTAATGAAGACAGTAGAAGAGAGACCTAGAGATTGAGGCTTAA 1385  
 Db |||||  
 QY 1393 AGAACGAATGTTGCTCATGAGACAGTGGAGGAAAGAACTTAGAGATTGAGAGGCTTAA 1452  
 Db |||||  
 QY 1386 GACGAACAAAGAACTGAGGAGCAAGATGTTGCGCCAGAGGCTGAGGA 1436  
 Db |||||  
 QY 1453 GATGAGCAAAAGAACTGGAAGCAAGGTTGTTGCGCCAGAGGACCCGGA 1503  
 Db |||||

RESULT 6  
 AL545712 1124 bp mRNA linear EST 25-MAR-2004  
 LOCUS  
 DEFINITION AL545712 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 Clone CSODI009YJ01 5-PRIME, mRNA sequence.

ACCESSION AL545712  
 VERSION  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homniidae; Homo.  
 1 (bases 1 to 1124)  
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:31267547.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 7766.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?c=CSODI009CE01QI1&c=7766.r.

#### FEATURES

Location/Qualifiers  
 1. 1124  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODI009YJ01"  
 /tissue type="PLACENTA COT 25-NORMALIZED"  
 /clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

#### ORIGIN

Query Match 68.8%; Score 1007; DB 1; Length 1124;  
 Best Local Similarity 96.9%; Pred. No. 1.7e-257;  
 Matches 101; Conservative 25; Mismatches 5; Indels 2; Gaps 2;  
 QY 4 GGTGCTGTGGCTAAGCAAGATTGGAGTACTGCTGCCACCTCCAGCTCGGCTAAGG 63  
 Db |||||  
 QY 59 GGTGCTGTGGCTAAGCAAGATTGGAGTACTGCTGCCACCTCCAGCTCGGCTAAGG 118  
 Db |||||  
 QY 64 GTGCTGTGGACTGCGGCCATTGTTGGATGGAAACAGCGGAGCAAGTATCCCCCTGT 123  
 Db |||||  
 QY 119 GTGCTGTGGACTGCGGCCATTGTTGGATGGAAACAGCGGAGCAAGTATCCCCCTGT 178  
 Db |||||  
 QY 124 GTGCGGGCATGGACAGCTGCTCTAGAGATGCTTAACCTGGAGAACCCACAGAGACT 183  
 Db |||||  
 QY 179 GTGCGGGCATGGACAGCTGCTCTAGAGATGCTTAACCTGGAGAACCCACAGAGACT 238  
 Db |||||  
 QY 184 CTCAAAATACAGTTTGTATGCTCTTATGGGGAGAGAGTACTACAGAGGACATCTATGCA 243  
 Db |||||  
 QY 239 CTCAAAATACAGTTTGTATGCTCTTATGGGGAGAGAGTACTACAGAGGACATCTATGCA 298  
 Db |||||  
 QY 244 GGTTCAGTCAGGCCATCTTAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTCTTGC 303  
 Db |||||  
 QY 299 GGTTCAGTCAGGCCATCTTAAGGCACTTGTGGAAGGCGAGATGTCAGTGTCTTGC 358  
 Db |||||  
 QY 304 TATGGACCCACAGAGCTGGGAGAGCACACATGCTGGGCGCCAGAGCAACCTGGG 363  
 Db |||||  
 QY 359 TATGGACCCACAGAGCTGGGAGAGCACACATGCTGGGCGCCAGAGCAACCTGGG 418  
 Db |||||  
 QY 364 GTGATCCCGCGGCTCTCATGGACCTCTCATGGAGAGGAGGAGGTCGCGAGGGC 423  
 Db |||||  
 QY 419 GKATCCCGCGGCTCTCATGGACCTCTCATGGAGAGGAGGAGGTCGCGAGGGG 478  
 Db |||||  
 QY 424 CGGCGATGGGCCCTTTCTGTACATGCTTTACTAGAGATCTTACAGGAGAAAGTATTA 483  
 Db |||||  
 QY 479 CGGCGATGGGCCCTTTCTGTACATGCTTTACTAGAGATCTTACAGGAGAAAGKATTA 538  
 Db |||||  
 QY 484 GACCTCTGACCTCTGCTGGGAGAGCTGATCCGAGAGACTCCCGGGGAATATC 543  
 Db |||||  
 QY 539 GACCTCTGACCTCTGCTGGGAGAGCTGATCCGAGAGACTCCCGGGGATATC 598  
 Db |||||  
 QY 544 CTGATTCGCGGCTCTCTCCAGAGACCCATCAGTAGCTTTTGTGATTTGAGCGGCATTC 603  
 Db |||||  
 QY 599 CTGATTCGCGGCTCTCTCCAGAGACCCATCAGTAGCTTTTGTGATTTGAGCGGCATTC 658  
 Db |||||  
 QY 604 CTGCGAGCCAGTCAAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC 663  
 Db |||||  
 QY 659 CTGCGAGCCAGTCAAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC 718  
 Db |||||  
 QY 664 AGTCATGCTGTGCTCTGCTGAGTGGACAGCGGGAACGTTTGGGCCCATTTGCGCAG 723  
 Db |||||  
 QY 719 AGTCATGCTGTGCTCTGCTGAGTGGACAGCGGGAACGTTTGGGCCCATTTGCGCAG 778  
 Db |||||  
 QY 724 CGAGAGGGAATACTTACCTGATTGACTTGGCTGGGTGAGAGGACACACCGCGCACAGGC 783  
 Db |||||











Db 309 GGTTCAGTGCAGCCATCCTAAGGCACTTGTCTGGAGGGCAGATGCCAGTGTCTTGCC 368  
Qy 304 TATGACCCACAGGAGCTGGGAAGACGACACATGTCTGGGAGCCACAGAGCAACCTGGG 363  
Db 369 TATGACCCACAGGAGCTGGGAAGACGACACATGTCTGGGAGCCACAGAGCAACCTGGG 428  
Qy 364 GTGATCCCGGGCTCTCATGACCTCTGTGACATCTGTGAGCTCAACAGGAGGAGGTGCCAGGGC 423  
Db 429 GTGATCCCGGGCTCTCATGACCTCTGTGACATCTGTGAGCTCAACAGGAGGAGGTGCCAGGGC 488  
Qy 424 CGGCATCGGGCCCTTCTGTGACCATGTCTTACCTAGAGATCTACACAGGAGAGGTATTA 483  
Db 489 CGGCATCGGGCCCTTCTGTGACCATGTCTTACCTAGAGATCTACACAGGAGAGGTATTA 548  
Qy 484 GACCTCTGGAGCCCTTCTGGGAGACCTGTGTAATCCGAGAGAGTGTCCGGGGGAATATC 543  
Db 549 GACCTCTGGAGCCCTTCTGGGAGACCTGTGTAATCCGAGAGAGTGTCCGGGGGAATATC 608  
Qy 544 CTGATTCGGGTCTCTCCAGAGCCATCATGAGTCTGAGTGTGTAATTTGAGCGGACATTC 603  
Db 609 CTGATTCGGGTCTCTCCAGAGCCATCATGAGTCTGAGTGTGTAATTTGAGCGGACATTC 668  
Qy 604 CTGCAGCCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCGGC 663  
Db 669 CTGCAGCCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAAMWAGCGCTCTCCGGC 728  
Qy 664 AGTCATGCTGTCTCTGGTCAAGGTGAGCAGCGGGAACCTTTGGGCCCAATTTTCGCAG 723  
Db 729 AGTCATGCTGTCTCTGGTCAAGGTGAGCAGCGGGAACCTTTGGGCCCAATTTTCGCAG 788  
Qy 724 CGAGAGGGAATCTTACCTGATTTGACTTGGCTGGTCAAGGAGCAACCGGCGCACAGGC 783  
Db 789 CGAGAGGGAATCTTACCTGATTTGACTTGGCTGGTCAAGGAGCAACCGGCGCACAGGC 848  
Qy 784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTGTCTCGGC 843  
Db 849 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTGTCTCGGC 908  
Qy 844 AAAGTGTAGTGTGGTGAATCAGGGCTCCCTGCTGACCTTATCGGAGCAGCAAGCTC 903  
Db 909 AAAGTGTAGTGTGGTGAATCAGGGCTCCCTGCTGACCTTATCGGAGCAGCAAGCTC 968  
Qy 904 ACTGCTTATGAGGACTCTCTGGTGGTCTCAGCCACAGTATCTTATTTGCCAAT 963  
Db 969 CTGCGCTATTGCA-GRCTCTCTGGTGGTCAACCCMAGTAT---CVTATTGCAATTTG 1024  
Qy 964 GCCCTGAGAGACCTTCTA 983  
Db 1025 CCCTKARAGCTTTTHCTA 1044

RESULT 11  
AL518361 1016 bp mRNA linear EST 23-MAR-2004  
LOCUS  
DEFINITION CSODA009YD01 5-PRIME, mRNA sequence.  
ACCESSION AL518361  
VERSION AL518361.3 GI:45654904  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1016)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 13, 2001 this sequence version replaced gi:30536600.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
This sequence belongs to sequence cluster 7766.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CSODA009CB01QPI&c=7766.r.  
Location/Qualifiers  
1. 1016  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODA009YD01"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_lib="Homo sapiens NEUROBLASTOMA"  
/note="vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match 61.3%; Score 897; DB 1; Length 1016;  
Best Local Similarity 98.5%; Pred. No. 4.1e-228;  
Matches 923; Conservative 3; Mismatches 8; Indels 3; Gaps 2;  
Qy 4 GGTGCTGTGGCTTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGGTAAGG 63  
Db 9 GGTGCTGTGGCTTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGGTAAGG 68  
Qy 64 GTGCTGTGGCTTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGGTAAGG 123  
Db 69 GTGCTGTGGCTTAAGCAAGATTGGAGTACTCGTCCACCTCCAGCTCGGTAAGG 128  
Qy 124 GTGCGGGCATGAGCAGCTGCTCTAGAGATTCCTAATCGAGAGAACCAACAGAGACT 183  
Db 129 GTGCGGGCATGAGCAGCTGCTCTAGAGATTCCTAATCGAGAGAACCAACAGAGACT 188  
Qy 184 CTCAATACCAAGTTTGTGATGCTTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 243  
Db 189 CTCAATACCAAGTTTGTGATGCTTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 248  
Qy 244 GGTTCAGTCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGATGCCAGTGTCTTGCC 303  
Db 249 GGTTCAGTCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGATGCCAGTGTCTTGCC 308  
Qy 304 TATGACCCACAGGAGCTGGAGAGACCAACATGTCTGGGAGCCAGAGCAACCTGGG 363  
Db 309 TATGACCCACAGGAGCTGGAGAGACCAACATGTCTGGGAGCCAGAGCAACCTGGG 368  
Qy 364 GTGATCCCGGGCTCTCATGGACCTCCTCGAGCTCAAGGGAGGAGGTGCCAGGGC 423  
Db 369 GTGATCCCGGGCTCTCATGGACCTCCTCGAGCTCAAGGGAGGAGGTGCCAGGGC 428  
Qy 424 CGGCATCGGGCCCTTCTGTGACCATGTCTTACCTAGAGATCTACACAGGAGAGGTATTA 483  
Db 429 CGGCATCGGGCCCTTCTGTGACCATGTCTTACCTAGAGATCTACACAGGAGAGGTATTA 488  
Qy 484 GACCTCTGGAGCCCTTCTGGGAGACCTGTGTAATCCGAGAGACTGCCGGGGGAATATC 543  
Db 489 GACCTCTGGAGCCCTTCTGGGAGACCTGTGTAATCCGAGAGACTGCCGGGGGAATATC 548  
Qy 544 CTGATTCGGGTCTCTCCAGAGCCATCAGTAGTCTTCTGATTTTGGCGGACATTC 603  
Db 549 CTGATTCGGGTCTCTCCAGAGCCATCAGTAGTCTTCTGATTTTGGCGGACATTC 608  
Qy 604 CTGCAGCCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGGC 663  
Db 609 CTGCAGCCAGTCCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGGC 668  
Qy 664 AGTCATGCTGTCTCTGGTCAAGGTGAGCAGCGGGAACCTTTGGGCCCAATTTTCGCAG 723  
Db 669 AGTCATGCTGTCTCTGGTCAAGGTGAGCAGCGGGAACCTTTGGGCCCAATTTTCGCAG 728

QY	724	CGAGAGGGAATACTCTACCTG-ATTGACTTGGCTGGGTACAGGACAACCGCGCACAGG	782	Db	61	AGCCAGAGCAACTGGGTGATCCCGGGCTCTCATGACCTCTGACGCTCAACAGG	120
Db	729	CGAGAGGGAATACTCTACCTGATTTGACTTGGCTGGGTACAGGACAACCGCGCACAGG	788	QY	406	GAGAGGGTCCCGAGGCGCGGCATATGGGCCCTTTTGTGTACCATGTCTTACCTAGAGATC	465
QY	783	CAACAAGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCTGGG	842	Db	121	GAGAGGGTCCCGAGGCGCGGCATATGGGCCCTTTTGTGTACCATGTCTTACCTAGAGATC	180
Db	789	CAACAAGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCTGGG	848	QY	466	TACCAGAGAGGTATTTAGACCTCTCCGACCTCTCTCGGGAGACCTGGTAAATCCGAGAA	525
QY	843	CAAAAGTGTAGTGGCTGAATCAGGCGCTCCCTGTGTACCTTATCGGACAGCAAGCT	902	Db	181	TACCAGAGAGGTATTTAGACCTCTCTCGACCTCTCTCGGGAGACCTGGTAAATCCGAGAA	240
Db	849	CAAAAGTGTAGTGGCTGAATCAGGCGCTCCCTGTGTACCTTATCGGACAGCAAA--C	906	QY	526	GACTGCGGGGGGAATATCTCTGATTTCCGGGTCTCTCCAGAAAGCCATCAGTAGCTTTGCT	585
QY	903	CACTCCCTATTGACGAGACTCTCTGGGTGGCTCAGCC	939	Db	241	GACTGCGGGGGGAATATCTCTGATTTCCGGGTCTCTCCAGAAAGCCATCAGTAGCTTTGCT	300
Db	907	CCACTCGCTATTGACGAGACTCTCTGGGTGGCTCAGCC	943	QY	586	GATTTGAGCGGCACTTCTCGCAGCCAGTTCGAATCGGACTGTAGGAGCACCACCGGCTC	645
RESULT 12				Db	301	GATTTGAGCGGCACTTCTCGCAGCCAGTTCGAATCGGACTGTAGGAGCACCACCGGCTC	360
BQ051730				QY	646	AACAGCGCTCTCCCGCAGTCACTGTGTCTCTCGTCAAGGTGGACCCAGCGGAAACGT	705
LOCUS				Db	361	AACAGCGCTCTCTCCCGCAGTCACTGTGTCTCTCGTCAAGGTGGACCCAGCGGAAACGT	420
DEFINITION				QY	706	TTGGCCCCCATTTCCGACGAGAGGGAATACTCTACCTGATTGACTTTGGCTGGGTACAG	765
ACCESSION				Db	421	TTGGCCCCCATTTCCGACGAGAGGGAATACTCTACCTGATTGACTTTGGCTGGGTACAG	480
VERSION				QY	766	GACAAACCGGCGCACAGGCAACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACC	825
KEYWORDS				Db	481	GACAAACCGGCGCACAGGCAACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACC	540
SOURCE				QY	826	TCCCTGTTTGTCTCTGGGCAAAAGTGTAGATCGCTGTAATCAGGGCTCTCCCTCGTGTACT	885
ORGANISM				Db	541	TCCCTGTTTGTCTCTGGGCAAAAGTGTAGATCGCTGTAATCAGGGCTCTCCCTCGTGTACT	600
REFERENCE				QY	886	TATCGGACAGCAAGCTCACTCGCTTATTCAGGACTCTCTGGGTGGCTCAGCCACACAGT	945
AUTHORS				Db	601	TATCGGACAGCAAGCTCACTCGCTTATTCAGGACTCTCTGGGTGGCTCAGCCACACAGT	660
TITLE				QY	946	ATCCCTTATTTGCCAACAATTCGCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGCACTC	1005
JOURNAL				Db	661	ATCCCTTATTTGCCAACAATTCGCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGCACTC	720
COMMENT				QY	1006	AAC-TTTGCTGCCAGGTCCAAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGCA	1064
FEATURES				Db	721	AACTTTTGTCTGCCAGGTCCAAGAGGTGATCAATCGGCTTTTA-CAATGAGAGCTGCA	779
source				QY	1065	GCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACCCAGAGGC	1124
1. .1034				Db	780	GCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACCCAGGC	839
/organism="Homo sapiens"				QY	1125	AAAGAGAGCCC--GAGGCCCTCAGGAAGAGGAGATTGGAG--CCCTGAGCCCATGGCA--	1179
/db_xref="taxon:9606"				Db	840	AAAGAGAGCCCAGAGGCCCTCAGGAAGAGGAGATTGGGAGCCCTTAGCCCATGGCAAG	899
/clones="IMAGE:5932716"				QY	1180	GCTCCAGCTCTGCTTCCAGAAAACCTCAG--CCCTCAGAGAGCTTAAGCAGCATGGACCC	1238
/tissue type="natural killer cells, cell line"				Db	900	CTCCAGCCCTCTGCTTCCAGAAAACCTCAGCCCTTACAGAGCTTAGCAGCATGGACCC	959
/lab host="DH10B (phage-resistant)"				QY	1239	GGCCA-TGCTGGAGCGCTCTCAGCTT--GGACCGTCTGCTT--GCCTTCCAGGGGGAGC	1293
/clone lib="NIH MGC 106"				Db	960	GGCCACTGTGTGAGCGCCCTCTTCAGCTTTGGAGCCCTGCTGCTTTGCCTCCAGGGGGAGC	1019
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				QY	1294	CAGGGGGCCCCCTCTG	1308
				Db	1020	CAGGGGGCCCCCTCGG	1034
Query Match							
Best Local Similarity							
Matches							
999; Conservative							
0; Mismatches							
23; Indels							
13; Gaps							
9;							
QY	286	AATGCCAGTGTCTTGGCTTATGGACCCACAGAGCTGGGAGAGCGCACACAAATGCTGGGC	345				
Db	1	AATGCCAGTGTCTTGGCTTATGGACCCACAGAGCTGGGAGAGCGCACACAAATGCTGGGC	60				
QY	346	AGCCACAGACCACTGGGGTGATCCCGGGGCTCTCATGGACCTCTCTGACGCTCACAGG	405				

RESULT 13  
BX363358  
LOCUS  
DEFINITION  
BX363358 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0K012YE23 5-PRIME, mRNA sequence.  
959 bp mRNA linear EST 08-APR-2004

ACCESSION BX363358  
VERSION BX363358.2 GI:46287820  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 959)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 5, 2003 this sequence version replaced gi:30368690.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7766.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0DK012AC120P1&c=7766.r.  
FEATURES  
source  
1..959  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK012YE23"  
/cell\_type="HELA CELLS"  
/cell\_line="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
ORIGIN  
Query Match 60.2%; Score 881; DB 5; Length 959;  
Best Local Similarity 100.0%; Pred. No. 7.5e-224;  
Matches 881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 GGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAAG 63  
Db 79 GGTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAAG 138  
Qy 64 GTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAAG 123  
Db 139 GTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGGTAAAG 198  
Qy 124 GTGGGGGCGATGGACAGCTCTCTAGAGATTCTTAAGTGGAGGAGGAGGAGGAGGAG 183  
Db 199 GTGGGGGCGATGGACAGCTCTCTAGAGATTCTTAAGTGGAGGAGGAGGAGGAGGAG 258  
Qy 184 CTCAATATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 243  
Db 259 CTCAATATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 318  
Qy 244 GGTTCAGTCAGGCCCATCTTAAGGCACTTGTGTGAAGGCGAGAGTGCAGTGTCTTGCC 303  
Db 319 GGTTCAGTCAGGCCCATCTTAAGGCACTTGTGTGAAGGCGAGAGTGCAGTGTCTTGCC 378  
Qy 304 TATGGACCCACAGAGCTGGGAAGCAGCAGCATATGCTGGGAGGAGGAGGAGGAGGAGG 363  
Db 379 TATGGACCCACAGAGCTGGGAAGCAGCAGCATATGCTGGGAGGAGGAGGAGGAGGAGG 438  
Qy 364 GTGATCCCGGGGCTCTCATGGACTCTCTGAGCTCTCAAGGGAGGAGGAGGAGGAGGAG 423  
Db 439 GTGATCCCGGGGCTCTCATGGACTCTCTGAGCTCTCAAGGGAGGAGGAGGAGGAGGAG 498  
Qy 424 CGGCCATGGGCGCTTCTGTGCACCATGTCTTACTAGAGATCTTACCGAGGAGAGGTATTA 483

Db 499 CGGCCATGGGCGCTTCTGTGCACCATGTCTTACCTAGAGATCTTACCGAGAGAGGTATTA 558  
Qy 484 GACCTCTGGACCCCTGCTTCGGGAGAGACTGTGTAATCCGAGAGACTGCGGGGGGAATATC 543  
Db 559 GACCTCTGGACCCCTGCTTCGGGAGAGACTGTGTAATCCGAGAGACTGCGGGGGGAATATC 618  
Qy 544 CTGATTCGGGCTCTCTCCAGAGCCCATCATAGTAGCTTTGCTGATTTTGGAGCGGCACATC 603  
Db 619 CTGATTCGGGCTCTCTCCAGAGCCCATCATAGTAGCTTTGCTGATTTTGGAGCGGCACATC 678  
Qy 604 CTGCCAGCAGTCGAATTCGAGCTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCG 663  
Db 679 CTGCCAGCAGTCGAATTCGAGCTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCG 738  
Qy 664 AGTCATGCTGTGCTCTCGCTCAAGGTGGACAGCGGGAACGTTTGGCCCAATTTCCGCGAG 723  
Db 739 AGTCATGCTGTGCTCTCGCTCAAGGTGGACAGCGGGAACGTTTGGCCCAATTTCCGCGAG 798  
Qy 724 CGAGAGGAGAAACTCTTACCTGATTTGACTTGGCTCGGTCAGAGGACAAACCGCGCACAGGC 783  
Db 799 CGAGAGGAGAAACTCTTACCTGATTTGACTTGGCTCGGTCAGAGGACAAACCGCGCACAGGC 858  
Qy 784 AACAGAGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACTCTCTGTTGCTCTGGGC 843  
Db 859 AACAGAGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACTCTCTGTTGCTCTGGGC 918  
Qy 844 AAAGTGTAGATGCGCTGAATTCAGGCGCTCCCTCGTGTACC 884  
Db 919 AAAGTGTAGATGCGCTGAATTCAGGCGCTCCCTCGTGTACC 959  
RESULT 14  
LOCUS B0053671  
DEFINITION B0053671 971 bp mRNA linear EST 29-MAR-2002  
5', mRNA sequence.  
ACCESSION B0053671  
VERSION B0053671.1 GI:19813011  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 971)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE NIH-MGC  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM2123 row: g column: 07  
High quality sequence stop: 660.  
Location/Qualifiers  
1..971  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5935518"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 106"  
/notes="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and



```

Db      373  TATGACCCACAGAGCTGGGAAGCAGCACAAATGCTGGGACGCCAGAGCAACCTGGG 432
Qy      364  GTGATCCCGGGGCTCTCATGACCTCTCTGAGCTCAAGGGAGGAGGTGCCGAGGGC 423
Db      433  GTGATCCCGGGGCTCTCATGACCTCTCTGAGCTCAAGGGAGGAGGTGCCGAGGGC 492
Qy      424  CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACACAGGAGAGGTATTA 483
Db      493  CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACACAGGAGAGGTATTA 552
Qy      484  GACCTCTGTGACCTGTCTTGGGAGACTGTGTAATCCGAGAAGACTCCGCGGGGAATATC 543
Db      553  GACCTCTGTGACCTGTCTTGGGAGACTGTGTAATCCGAGAAGACTCCGCGGGGAATATC 612
Qy      544  CTGATTCGGGTCTCTCCAGAACCCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTC 603
Db      613  CTGATTCGGGTCTCTCCAGAACCCCATCAGTAGCTTTGCTGATTTTGAGCGGCATTC 672
Qy      604  CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGGTCAACCAAGCGCTCTCCCGC 663
Db      673  CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGGTCAACCAAGCGCTCTCCCGC 732
Qy      664  AGTCATGCTGTGCTCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGCCCAATTCGCCAG 723
Db      733  AGTCATGCTGTGCTCTCTGCTCAAGGTGAGCCAGCGGGAACGTTTGGCCCAATTCGCCAG 792
Qy      724  CGAGAGGGAAAACCTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGCACAGGC 783
Db      793  CGAGAGGGAAAACCTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGCACAGGC 852
Qy      784  AACAGGGCCCTTCGGCTAAAAGAGAGTGGAGGCCATCAACCTCCCTGTTTGTCTGGGC 843
Db      853  AACAGGGCCCTTCGGCTAAAAGAGAGTGGAGGCCATCAACCTCCCTGTTTGTCTGGGC 912
Qy      844  -AAAGTGTAGATCGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTAT 888
Db      913  AAAAGTGTAGATCGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTAT 958

```

Search completed: November 10, 2004, 18:18:12  
 Job time : 5093.37 secs

**This Page Blank (uspto)**

Query Match	100.0%	Score 1115;	DB 6;	Length 1115;
Best Local Similarity	100.0%	Prod. No. 2.7e-278;		
Matches 1115;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	ATGCAGCGCGGGCGGCTTCGACGCAGCAGAGGCGCGCAGAGATGGCGGAGCTTCAGCGG	60		
1	ATGCAGCGCGGGCGGCTTCGACGCAGCAGAGGCGCGCAGAGATGGCGGAGCTTCAGCGG	60		
61	CGGCGATCTCAGGAGCTGTGCTCGGCTTAAGCAAGATTGGAGCTACTCTGTCGTCCAC	120		
61	CGGCGATCTCAGGAGCTGTGCTCGGCTTAAGCAAGATTGGAGCTACTCTGTCGTCCAC	120		
121	CTCAGCTCGCGTTAAGGGTGGCTGTGCGCATCTGGCGCATTTGTGGATGGAACAGCGGGAG	180		
121	CTCAGCTCGCGTTAAGGGTGGCTGTGCGCATCTGGCGCATTTGTGGATGGAACAGCGGGAG	180		
181	CAAGTATCCCCCTGTGTGGGGGATGACACAGCTGCTCTTAGAGATTGCTAACTGGA	240		
181	CAAGTATCCCCCTGTGTGGGGGATGACACAGCTGCTCTTAGAGATTGCTAACTGGA	240		
241	GGAAACACAGGAGACTCTCAAAATACCAGTTTGCATGCTTCTATGGGGAGAGGACTCTC	300		

Db 241 GGRACCACGAGGACTCTCAATACCAAGTTGATGCTTCTATGGGAGGAGTACTC 300  
Qy 301 AGCAGGACATCTATGCAAGTTTCAGTGCAGGCCATCTTAAGCACTTCTGTAAGGGGAGG 360  
Db 301 AGCAGGACATCTATGCAAGTTTCAGTGCAGGCCATCTTAAGCACTTCTGTAAGGGGAGG 360  
Qy 361 ATGCGAGTGTCTGCTTATGAGCCACAGAGGCTGGGAAGAGCCACACATGCTGGGCA 420  
Db 361 ATGCGAGTGTCTGCTTATGAGCCACAGAGGCTGGGAAGAGCCACACATGCTGGGCA 420  
Qy 421 GCCCAGAGCAACCTGGGGTATATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAAGG 480  
Db 421 GCCCAGAGCAACCTGGGGTATATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAAGG 480  
Qy 481 AGGAGGTGCGGAGGCGGCGCATGGGCTTCTGTCATCAATGCTCTTACCTAGATCT 540  
Db 481 AGGAGGTGCGGAGGCGGCGCATGGGCTTCTGTCATCAATGCTCTTACCTAGATCT 540  
Qy 541 ACCAGGAGAGGTATATAGACCTCTGAGACCTGCTTCGGGAGACCTGCTAATCCGAGAAG 600  
Db 541 ACCAGGAGAGGTATATAGACCTCTGAGACCTGCTTCGGGAGACCTGCTAATCCGAGAAG 600  
Qy 601 ACTCGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660  
Db 601 ACTCGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660  
Qy 661 ATTTTGAAGCGCATCTTCTGACGCGAGTCAATCGAATCGAATCGAATCGAATCGAATCGA 720  
Db 661 ATTTTGAAGCGCATCTTCTGACGCGAGTCAATCGAATCGAATCGAATCGAATCGAATCGA 720  
Qy 721 ACCAGCGCTCTCCGCGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 780  
Db 721 ACCAGCGCTCTCCGCGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 780  
Qy 781 TGGCGCCATTTTGGCAGGAGGAAATCTACCTGATGACTTGGCTGGGTGAGGAGGAGGAGG 840  
Db 781 TGGCGCCATTTTGGCAGGAGGAAATCTACCTGATGACTTGGCTGGGTGAGGAGGAGGAGG 840  
Qy 841 ACAACCGCGCACAGGCAACAGGCGCTTCGGCTTAAAGAGAGTGGAGGCCATCAACACCT 900  
Db 841 ACAACCGCGCACAGGCAACAGGCGCTTCGGCTTAAAGAGAGTGGAGGCCATCAACACCT 900  
Qy 901 CCCTGTTTGTCTCGGCAAGTGTAGATGGCTGAATCAGGGCTCTCCCTGCTGACCTT 960  
Db 901 CCCTGTTTGTCTCGGCAAGTGTAGATGGCTGAATCAGGGCTCTCCCTGCTGACCTT 960  
Qy 961 ATCGGACAGCAAGCTCACTCGCCTATTCAGGAGCTCTCTGGGTGGGTGAGGAGGAGGAGG 1020  
Db 961 ATCGGACAGCAAGCTCACTCGCCTATTCAGGAGCTCTCTGGGTGGGTGAGGAGGAGGAGG 1020  
Qy 1021 TCCTTATTTGCCAACAATTGCGGCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
Db 1021 TCCTTATTTGCCAACAATTGCGGCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
Qy 1081 ACTTTGCTGCCAGGTCCAGGAGGTGATCAATTGA 1115  
Db 1081 ACTTTGCTGCCAGGTCCAGGAGGTGATCAATTGA 1115

RESULT 2  
AR210053  
LOCUS AR210053 1538 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 3 from patent US 6387644.  
ACCESSION AR210053  
VERSION AR210053.1 GI:21512183  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1538)  
AUTHORS Beraud, C.  
TITLE Motor proteins and methods for their use  
JOURNAL Patent: US 6387644-A 3 14-MAY-2002;

FEATURES  
source 1..1538  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 99.8%; Score 1112.4; DB 6; Length 1538;  
Best Local Similarity 99.9%; Pred. No. 1.3e-277;  
Matches 1113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATCCAGCCGCGGCGCTCGACCCAGAGAGGCGAGCGGAGATGGCGGAGCTTCAAGCGG 60  
Db 1 ATCCAGCCGCGGCGCTCGACCCAGAGAGGCGAGCGGAGATGGCGGAGCTTCAAGCGG 60  
Qy 61 CGCGGATCTCAGAGCTGCTGCGTGTGCGTGTAGCAAGATTGAGACTACTCTGCTCCAC 120  
Db 61 CGCGGATCTCAGAGCTGCTGCGTGTGCGTGTAGCAAGATTGAGACTACTCTGCTCCAC 120  
Qy 121 CTCAGCTCGCGTAAGGCTGCTGCGACTGCGGCCATTTTGGATGGAACAGCGGAG 180  
Db 121 CTCAGCTCGCGTAAGGCTGCTGCGACTGCGGCCATTTTGGATGGAACAGCGGAG 180  
Qy 181 CAAGTATCCCTCTGTGCGGGGCAATGAGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240  
Db 181 CAAGTATCCCTCTGTGCGGGGCAATGAGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240  
Qy 241 GGAACACAGGAGACTCTCAATATACCAAGTTTGTATGCTTCTATGGGAGGAGTACTC 300  
Db 241 GGAACACAGGAGACTCTCAATATACCAAGTTTGTATGCTTCTATGGGAGGAGTACTC 300  
Qy 301 AGCAGGACATCTATGCAAGTTTCAGTGCAGGCCATCTTAAGCACTTCTGTAAGGGGAGG 360  
Db 301 AGCAGGACATCTATGCAAGTTTCAGTGCAGGCCATCTTAAGCACTTCTGTAAGGGGAGG 360  
Qy 361 ATGCGAGTGTCTGCTTATGAGCCACAGAGGCTGGGAAGAGCCACACATGCTGGGCA 420  
Db 361 ATGCGAGTGTCTGCTTATGAGCCACAGAGGCTGGGAAGAGCCACACATGCTGGGCA 420  
Qy 421 GCCCAGAGCAACCTGGGGTATATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAAGG 480  
Db 421 GCCCAGAGCAACCTGGGGTATATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAACAAGG 480  
Qy 481 AGGAGGTGCGGAGGCGGCGCATGGGCTTCTGTCACCATGCTTCTTACCTAGAGATCT 540  
Db 481 AGGAGGTGCGGAGGCGGCGCATGGGCTTCTGTCACCATGCTTCTTACCTAGAGATCT 540  
Qy 541 ACCAGGAGAGGTATTAGACTCTCTGAGACCTCTTCCGAGAGCTTGGTAAATCCGAGAAG 600  
Db 541 ACCAGGAGAGGTATTAGACTCTCTGAGACCTCTTCCGAGAGCTTGGTAAATCCGAGAAG 600  
Qy 601 ACTCGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660  
Db 601 ACTCGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660  
Qy 661 ATTTTGAAGCGCATCTTCTGCGGAGGCTTCTTACCTGATGACTTGGCTGGGTGAGGAGG 720  
Db 661 ATTTTGAAGCGCATCTTCTGCGGAGGCTTCTTACCTGATGACTTGGCTGGGTGAGGAGG 720  
Qy 721 ACCAGCGCTCTCCGCGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 780  
Db 721 ACCAGCGCTCTCCGCGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 780  
Qy 781 TGGCGCCATTTTGGCAGGAGGAAATCTACCTGATGACTTGGCTGGGTGAGGAGGAGGAGG 840  
Db 781 TGGCGCCATTTTGGCAGGAGGAAATCTACCTGATGACTTGGCTGGGTGAGGAGGAGGAGG 840  
Qy 841 ACAACCGCGCACAGGCAACAGGCGCTTCGGCTTAAAGAGAGTGGAGGCCATCAACACCT 900  
Db 841 ACAACCGCGCACAGGCAACAGGCGCTTCGGCTTAAAGAGAGTGGAGGCCATCAACACCT 900  
Qy 901 CCCTGTTTGTCTCGGCAAGTGTAGATGGCTGAATCAGGGCTCTCCCTGCTGACCTT 960  
Db 901 CCCTGTTTGTCTCGGCAAGTGTAGATGGCTGAATCAGGGCTCTCCCTGCTGACCTT 960



QY 961 ATCGGACAGCAAGCTCACTCGCTATTGACGAGCTCTCTGGTGGCTCAGGCCACAGTA 1020  
 DB 961 ATCGGACAGCAAGCTCACTCGCTATTGACGAGCTCTCTGGTGGCTCAGGCCACAGTA 1020  
 QY 1021 TCCATTATGCCAACATTGCGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGCACTCA 1080  
 DB 1021 TCCATTATGCCAACATTGCGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGCACTCA 1080  
 QY 1081 ACTTCTGCTCGAGTCCAGGAGGATGATCAATTG 1114  
 DB 1081 ACTTCTGCTCGAGTCCAGGAGGATGATCAATCG 1114  
 RESULT 3  
 LOCUS BT007259 1998 bp mRNA linear PRI 13-MAY-2003  
 DEFINITION Homo sapiens kinesin-like 4 mRNA, complete cds.  
 ACCESSION BT007259  
 VERSION BT007259.1 GI:30583356  
 KEYWORDS FLI CDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1998)  
 Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
 Phelan,M. and Farmer,A.  
 Cloning of human full-length cDNAs in BD Creator(TM) System Donor  
 vector  
 Unpublished  
 2 (bases 1 to 1998)  
 Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
 Phelan,M. and Farmer,A.  
 Direct Submission  
 Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
 Circle, Palo Alto, CA 94303, USA  
 This cDNA clone is a part of a collection of human full length  
 expression clones generated by BD Biosciences Clontech and the  
 Harvard Institute of Proteomics. Each cDNA has been cloned in two  
 forms: with and without stop-codon (to allow fusion with C-terminal  
 tag). The cDNA has been directionally cloned using BD In-Fusion(TM)  
 cloning system between the Sali and HindIII sites of the pDNR-DUAL  
 vector. Additional sequences in the clone: 'ACC' after Sali site  
 and before 'ATG' to provide Kozak consensus sequence; 'GG' after  
 last codon and before HindIII site to maintain reading frame.  
 Clone distribution: http://bioinfo.clontech.com/orfclones.  
 Location/Qualifiers  
 1..1998  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="GH00245X1.0"  
 /clone\_lib="BD Creator(TM) CDS Library derived from MGC  
 collection"  
 /lab\_host="DH5alpha Ti resistant"  
 /note="Vector: pDNR-Dual"  
 1..1998  
 /codon\_start=1  
 /product="kinesin-like 4"  
 /protein\_id="AAP35923.1"  
 /db\_xref="GI:30583356"  
 /translation="MAAGSGTQRRREMAAASAAISGAGRCRLSKI GATRRPPRV  
 RVNLRFPVDTAGASPPCVRGMDSCSLEIANWRNHQETLKYQDFAYGERQQD  
 IYAGSVQPIRLHLLGQNASVLYGPTGAGTKTMLGSPQPGVIPRALMDLLOITRE  
 EGAERPNALSVTMSYLBIEYOKVLDLPASGLDIVREDCRGNLII PGLSKPTSSF  
 ADPERHLPASNRVTGATRLNRSRSHAVLLVKVDQRERLAPROREGKLYLIDLA  
 GSDNRRITGNKRLKESGINTSLFVKGVDALNQLRPVPRVDRSKLRLQLDLSLG  
 GSASHILIANAPERFVLDYTVSALNFAARSKVINRPTNIESLQHPALVPKLSQKE  
 LLGPPEAKRGRPEEEI GSPFMAAPASQKLSPLQKLSMDPAMLERLLSLDRLL  
 ASQSGQAPLLSTPKRERWLMKTVEEKDLERLTKQKELEAKMLQAEBKENHC

PTMLRPLSHRTVTGAKPLKKA VVMPQLIQEOAASPNAIEHILKNKRKRKLSDAL  
 BEBAEDCWELQISPPELLAHGROKILDLNREGSARDLSRSLORIPKPKAQLVGNREL  
 HGPPFQVEDLERVEGITGQMESFLKANILGLAAGORCGAS"

Query Match 99.3%; Score 1107.4; DB 9; Length 1998;  
 Best Local Similarity 99.9%; Pred. No. 2.6e-276;  
 Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

QY 6 GCGCGGGGGGCTCGACGACGAGAGGCGACCGGAGATGGCGGACGCTTCAGCGGGGGG 65  
 DB 4 GCGCGGGGGGCTCGACGACGAGAGGCGACCGGAGATGGCGGACGCTTCAGCGGGGGG 63  
 QY 66 ATCTCAGGAGCTGGTGGCTGCTGCGCTAAGCAAGATTGGAGCTACTCGTGCACCTCCA 125  
 DB 64 ATCTCAGGAGCTGGTGGCTGCTGCGCTAAGCAAGATTGGAGCTACTCGTGCACCTCCA 123  
 QY 126 GCTCGCGTAAGGGTGGCTGTCGACTCGCGGCCATTGTGGATGGAAACAGCGGGAGCAAGT 185  
 DB 124 GCTCGCGTAAGGGTGGCTGTCGACTCGCGGCCATTGTGGATGGAAACAGCGGGAGCAAGT 183  
 QY 186 GATCCCGCTGCTGGGGGCGATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAAC 245  
 DB 184 GATCCCGCTGCTGGGGGCGATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAAC 243  
 QY 246 CACCAAGGAGACTCTCAAAATACCAATTGATGCTTCTTAGGGGAGAGGAGTACTCAGCAG 305  
 DB 244 CACCAAGGAGACTCTCAAAATACCAATTGATGCTTCTTAGGGGAGAGGAGTACTCAGCAG 303  
 QY 306 GACATCTATGCAAGTTTCAAGTGCAGCGCCCATCTTAAGGCATCTTCTTGGAAAGGAGCAATGCC 365  
 DB 304 GACATCTATGCAAGTTTCAAGTGCAGCGCCCATCTTAAGGCATCTTCTTGGAAAGGAGCAATGCC 363  
 QY 366 AGTGTGCTTGGCTATGGACCCACAGGAGCTGGGAAAGACACACAATGCTGGGCGACCCCA 425  
 DB 364 AGTGTGCTTGGCTATGGACCCACAGGAGCTGGGAAAGACACACAATGCTGGGCGACCCCA 423  
 QY 426 GAGCAACCTGGGTGATCCCGGGGCTCTCATGGACCTCTCGACGCTCAAGAGGAGGAG 485  
 DB 424 GAGCAACCTGGGTGATCCCGGGGCTCTCATGGACCTCTCGACGCTCAAGAGGAGGAG 483  
 QY 486 GGTCCGAGGGCGCGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACACAG 545  
 DB 484 GGTCCGAGGGCGCGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACACAG 543  
 QY 546 GAGAAGGTATTAGACTCTCTGACCTCTGCGGAGACCTGGTAAATCGGAGAAAGACTGC 605  
 DB 544 GAGAAGGTATTAGACTCTCTGACCTCTGCGGAGACCTGGTAAATCGGAGAAAGACTGC 603  
 QY 606 CGGGGGAATATCTGATTTCCGGGCTCTCTCCAGAGAGCCCATCAGTAGCTTTTCTGATTTT 665  
 DB 604 CGGGGGAATATCTGATTTCCGGGCTCTCTCCAGAGAGCCCATCAGTAGCTTTTCTGATTTT 663  
 QY 666 GAGCGGCACTTCTCTGCCACGACGATCGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 725  
 DB 664 GAGCGGCACTTCTCTGCCACGACGATCGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 723  
 QY 726 CGTCTCTCCGAGTCTGCTGCTCTCTGCTCAAGGTGGACCGGAGAACGTTTGGCC 785  
 DB 724 CGTCTCTCCGAGTCTGCTGCTCTCTGCTCAAGGTGGACCGGAGAACGTTTGGCC 783  
 QY 786 CCATTTCCCGACGAGAGGAGAAACTCTACCTGATTGATCTGGCTGGGTTCAGAGGACACAC 845  
 DB 784 CCATTTCCCGACGAGAGGAGAAACTCTACCTGATTGATCTGGCTGGGTTCAGAGGACACAC 843  
 QY 846 CGGCGCAGGCAACAAGGGGCTTCTGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTG 905  
 DB 844 CGGCGCAGGCAACAAGGGGCTTCTGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTG 903  
 QY 906 TTTGTCTGGGCAAGTGGTATAGCTCGCTGAATCAGGGGCTCCCTCGTGTACCTTATCGG 965  
 DB 904 TTTGTCTGGGCAAGTGGTATAGCTCGCTGAATCAGGGGCTCCCTCGTGTACCTTATCGG 963

QY 966 GACAGCAAGCTCACTCCGCTATTTCAGAGACTCTCTGGTGGCTCAGCCACAGTATCCCTT 1025  
 DB 964 GACAGCAAGCTCACTCCGCTATTTCAGAGACTCTCTGGTGGCTCAGCCACAGTATCCCTT 1023  
 QY 1026 ATTGCCAACATTGCCCCCTGAGAGAGCGTCTTACCTAGACACAGTCTCCGCACTCAACTTT 1085  
 DB 1024 ATTGCCAACATTGCCCCCTGAGAGAGCGTCTTACCTAGACACAGTCTCCGCACTCAACTTT 1083  
 QY 1086 GCTGCCAGGTCACAGGAGGTGATCAATTG 1114  
 DB 1084 GCTGCCAGGTCACAGGAGGTGATCAATCG 1112

BT007888 1998 bp mRNA linear SYN 13-MAY-2003  
 DEFINITION Synthetic construct Homo sapiens kinesin-like 4 mRNA, partial cds.  
 ACCESSION BT007888  
 VERSION BT007888.1 GI:30584614  
 KEYWORDS FLI CDNA.  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 1 (bases 1 to 1998)  
 Kalline,N., Chen,X., Rolf,A., Halleck,A., Hines,L., Eisenstein,S.,  
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,  
 Phelan,M. and Farmer,A.  
 Cloning of human full-length CDSs in BD Creator(TM) System Donor  
 vector  
 Unpublished  
 2 (bases 1 to 1998)  
 Kalline,N., Chen,X., Rolf,A., Halleck,A., Hines,L., Eisenstein,S.,  
 Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,  
 Phelan,M. and Farmer,A.  
 Direct Submission  
 Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
 Circle, Palo Alto, CA 94303, USA  
 This CDS clone is a part of a collection of human full length  
 expression clones generated by BD Biosciences Clontech and the  
 Harvard Institute of Proteomics. Each CDS has been cloned in two  
 forms: with and without stop-codon (to allow fusion with C-terminal  
 tag). The CDS has been directionally cloned using BD In-Fusion(TM)  
 cloning system between the SalI and HindIII sites of the pDNR-DUAL  
 vector. Additional sequences in the clone: 'ACC' after SalI site  
 and before 'Arg' to provide kozak consensus sequence; 'CG' after  
 last codon and before HindIII site to maintain reading frame.  
 Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES  
 source  
 1..1998  
 /organism="synthetic construct"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:32630"  
 /clone="GH00245L1.0"  
 /clone\_lib="BD Creator(TM) CDS Library derived from MGC  
 collection"  
 /lab\_host="DH5alpha Tl resistant"  
 /note="Vector: pDNR-Dual"  
 1..>1998  
 /note="Mutations: 1997:Stop->Leu"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Homo sapiens kinesin-like 4"  
 /protein\_id="AAP36560.1"  
 /db\_xref="GI:30584615"

CDS  
 /translation="WAAGSTQRRREMAAASAAISAGRCRLSKIGATRRPPRV  
 RVAVLRFPVGTGASDPPCVRGMDSCSLFIANRNHQETLKYOFDAFYGERSTOOD  
 IYAGVQPIRLHLEGQNASVAYGPTGAGKHTMLGSPQPGVIPALMDLLQLTRE  
 IEAGRPWALSIVLYQVKVLDLDPASGLVIREDCRGNILIPGLSQKPISSF  
 ADDEHFPLPASNRVTGATRLNRSRSHVLLVKVDORERLAPPROEGKLYLDILA  
 GSDNRPTGKGLRLKESGAINTSFLVKGKVDALNOGLPVPYRDSKLTLLDLSLG  
 GSANSLTIANPERFYLDTVSALNFAARKEVINRPTNESHLOPHALGPVKLSQKE  
 LLGPEAKRAGPPEEEIGSPPEPMAAPASQKUSPQKSSMDPAMLERLLSLDLL  
 ASQSGQAPLLSTPKRERMVLMKTVEEKDLIERLTKTKQKLEAKMLAQKAEKBNHC

PTMLRPLSHRTVTGAKPLKAVVMPLOLIOEOAASPNABIHILKNKRKRKLESLDAL  
 EPEBAEDCWELQISPELLAHGRQKILDLINEGSARDLSRLQIRGPKKQALVIGWREL  
 HGFPSQVEDLERVEGITGQMESFLKANILGLAAGORCGASL"

Query Match 99.3%; Score 1107.4; DB 12; Length 1998;  
 Best Local Similarity 99.9%; Pred. No. 2.6e-276;  
 Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

QY 6 GCCGCGGGCGGCTCGACGACGAGAGCGGCGAGATGGCGCAGCTTCAGCGGGCGG 65  
 DB 4 GCCGCGGGCGGCTCGACGACGAGAGCGGCGAGATGGCGCAGCTTCAGCGGGCGG 63  
 QY 66 ATCTCAGGAGCTGCTCGCTGCTGCTAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCCA 125  
 DB 64 ATCTCAGGAGCTGCTCGCTGCTGCTGCTAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCCA 123  
 QY 126 GCTCGCGTAAAGGGTGGCTGTGCGACTGCGGCGCAATTTGTGGATGGAACAGCGGAGCAAGT 185  
 DB 124 GCTCGGTAAAGGGTGGCTGTGCGACTGCGGCGCAATTTGTGGATGGAACAGCGGAGCAAGT 183  
 QY 186 GATCCCCCTGTGTGCGGGCATATGACAGCTGCTCTTAGAGATTGCTTAATCTGAGGAAC 245  
 DB 184 GATCCCCCTGTGTGCGGGCATATGACAGCTGCTCTTAGAGATTGCTTAATCTGAGGAAC 243  
 QY 246 CACCAGGAGACTCTCAAAATACCAGTTTGTATGCCCTTCTATGGGAGAGGAGTACTCAGCAG 305  
 DB 244 CACCAGGAGACTCTCAAAATACCAGTTTGTATGCCCTTCTATGGGAGAGGAGTACTCAGCAG 303  
 QY 306 GACATCTATGACAGTTTCAGTGCAGGCCCATCTTAAGGCACCTTGTGAAAGGGCAGATGCC 365  
 DB 304 GACATCTATGACAGTTTCAGTGCAGGCCCATCTTAAGGCACCTTGTGAAAGGGCAGATGCC 363  
 QY 366 AGTGTGCTTCCCTATGACCCACAGAGCTGGGAAGACGACACACAATGCTGGGCGAGCCCA 425  
 DB 364 AGTGTGCTTCCCTATGACCCACAGAGCTGGGAAGACGACACACAATGCTGGGCGAGCCCA 423  
 QY 426 GAGCAACCTGGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCTCAAGGGAGGAG 485  
 DB 424 GAGCAACCTGGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCTCAAGGGAGGAG 483  
 QY 486 GGTGCCGAGGGCGGCGCATATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACCAG 545  
 DB 484 GGTGCCGAGGGCGGCGCATATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACCAG 543  
 QY 546 GAGAGGATATTAGACCTCTCGGACCTGCTTCGGGAGACCTGTAATCCGAGAGAGACTGC 605  
 DB 544 GAGAGGATATTAGACCTCTCGGACCTGCTTCGGGAGACCTGTAATCCGAGAGAGACTGC 603  
 QY 606 CGGGGGAATATCTGATTCCGGGGTCTCTCCAGAGGCCCATCATGATAGCTTTGCTGATTTT 665  
 DB 604 CGGGGGAATATCTGATTCCGGGGTCTCTCCAGAGGCCCATCATGATAGCTTTGCTGATTTT 663  
 QY 666 GAGCGGCATCTCTTCGCGCAGCGCATCGAATTCGAGCTGTAGAGGCCACCGGCTCAACAG 725  
 DB 664 GAGCGGCATCTCTTCGCGCAGCGCATCGAATTCGAGCTGTAGAGGCCACCGGCTCAACAG 723  
 QY 726 CGCTCTCCCGCAGCTCATGCTGCTCTGCTCAAGGTGGACACGAGCGGGAAACGTTTGGCC 785  
 DB 724 CGCTCTCCCGCAGCTCATGCTGCTCTGCTCAAGGTGGACACGAGCGGGAAACGTTTGGCC 783  
 QY 786 CCATTTCCCGCAGCGAGAGGGAAACCTCTACCTGATTGCTTGGCTGGGTGAGAGGACAAC 845  
 DB 784 CCATTTCCCGCAGCGAGAGGGAAACCTCTACCTGATTGCTTGGCTGGGTGAGAGGACAAC 843  
 QY 846 CGGGCGCAGGCAACAGGGCCCTTCGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905  
 DB 844 CGGGCGCAGGCAACAGGGCCCTTCGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 903  
 QY 906 TTTGCTCTGGGCAAGTGTAGATGCCCTGAAATCAGGGCCCTCCCTCGTGTACCTTATCGG 965  
 DB 904 TTTGCTCTGGGCAAGTGTAGATGCCCTGAAATCAGGGCCCTCCCTCGTGTACCTTATCGG 963

```
QY 966 GACAGCAGCTACTCGCCTATTGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCCCTT 1025
|
|
|
Db 964 GACAGCAAGCTACTCGCCTATTGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCCCTT 1023
|
|
|
QY 1026 ATTGCCAACATTGCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
|
|
|
Db 1024 ATTGCCAACATTGCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1083
|
|
|
QY 1086 GTCGCCAGTCCAAAGGAGGTGATCAATTG 1114
|
|
|
Db 1084 GTCGCCAGTCCAAAGGAGGTGATCAATCG 1112
|
|
|
RESULT 5
AR304057 2097 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 34 from patent US 6544766.
ACCESSION AR304057
VERSION AR304057.1 GI:31692955
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 2097)
TITLE Beraud,C., Ohashi,C., Sakowicz,R., Vaisberg,E., Wood,K. and Yu,M.
JOURNAL Human kinesins and methods of producing and purifying human
FEATURES Patent: US 6544766-A 34 08-APR-2003;
Location/Qualifiers
source 1..2097
/molecule="unknown"
/organism="genomic DNA"

ORIGIN
Query Match 99.3%; Score 1107.4; DB 6; Length 2097;
Best Local Similarity 99.9%; Pred. No. 2.6e-276;
Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GCCCGGGCGGCTCGACGACGAGCGGACGCGAGATGGCGGAGCTTCAGCGCGGG 65
|
|
|
Db 28 GCCCGGGCGGCTCGACGACGAGCGGACGCGAGATGGCGGAGCTTCAGCGCGGG 87
|
|
|
QY 66 ATCTCAGAGCTGCTCGCTGCTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACTCA 125
|
|
|
Db 88 ATCTCAGAGCTGCTCGCTGCTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACTCA 147
|
|
|
QY 126 GCTCCGCTAAGGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 185
|
|
|
Db 148 GCTCCGCTAAGGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 207
|
|
|
QY 186 GATCCCCCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 245
|
|
|
Db 208 GATCCCCCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 267
|
|
|
QY 246 CACCAAGAGACTCTCAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 305
|
|
|
Db 268 CACCAAGAGACTCTCAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 327
|
|
|
QY 306 GACATCTATCAGGTTTCACTGAGCGCCATCTTAAGGCACTTGTGAGGAGGAGATGCC 365
|
|
|
Db 328 GACATCTATCAGGTTTCACTGAGCGCCATCTTAAGGCACTTGTGAGGAGGAGATGCC 387
|
|
|
QY 366 AGTGTGCTTGCTATGGACCCACAGAGCTGGAGAGCGACACAAATGCTGGGACGCCA 425
|
|
|
Db 388 AGTGTGCTTGCTATGGACCCACAGAGCTGGAGAGCGACACAAATGCTGGGACGCCA 447
|
|
|
QY 426 GAGCAACTTGGGTTGATCCCGGGCTCTCATGAGACCTCTGACCTCACAAGGAGGAG 485
|
|
|
Db 448 GAGCAACTTGGGTTGATCCCGGGCTCTCATGAGACCTCTGACCTCACAAGGAGGAG 507
|
|
|
QY 486 GGTGCGGAGGCGCGCCATGGCCCTTCTGTACCATGCTTACCTAGAGATCTACCA 545
|
|
|
Db 508 GGTGCGGAGGCGCGCCATGGCCCTTCTGTACCATGCTTACCTAGAGATCTACCA 567
|
|
|
```

```
QY 546 GAGAAAGTATTAGACCTCTCTGACCTCTCTCGGAGACCTGTTAATCCGAGAAAGTGC 605
|
|
|
Db 568 GAGAAAGTATTAGACCTCTCTGACCTCTCTCGGAGACCTGTTAATCCGAGAAAGTGC 627
|
|
|
QY 606 CGGGGGAATATCTCTGATTCGGGTCTCTCCAGAGACCCATCAGTAGACTTTGCTGATTTT 665
|
|
|
Db 628 CGGGGGAATATCTCTGATTCGGGTCTCTCCAGAGACCCATCAGTAGACTTTGCTGATTTT 687
|
|
|
QY 666 GAGCGGCACTTCTCTGCGCAGCTCGAATCGACTGTAGGAGCCACCCGGCTCAACCCAG 725
|
|
|
Db 688 GAGCGGCACTTCTCTGCGCAGCTCGAATCGACTGTAGGAGCCACCCGGCTCAACCCAG 747
|
|
|
QY 726 CGCTCTCTCCGCACTCATGCTGTCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCC 785
|
|
|
Db 748 CGCTCTCTCCGCACTCATGCTGTCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCC 807
|
|
|
QY 786 CCATTTGCGCCAGCGAGAGGAGAACTCTACCTGATGACTTGGCTGGGTGAGAGCAAC 845
|
|
|
Db 808 CCATTTGCGCCAGCGAGAGGAGAACTCTACCTGATGACTTGGCTGGGTGAGAGCAAC 867
|
|
|
QY 846 CGGCGCACAGCAACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
|
|
|
Db 868 CGGCGCACAGCAACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 927
|
|
|
QY 906 TTTGTCTCTGGGCAAAAGTGTAGATCGCTGTAATCAAGGGCTCCCTCGTGTACTTATCGG 965
|
|
|
Db 928 TTTGTCTCTGGGCAAAAGTGTAGATCGCTGTAATCAAGGGCTCCCTCGTGTACTTATCGG 987
|
|
|
QY 966 GACAAAGCTCACTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCCACAGTATCTCTT 1025
|
|
|
Db 988 GACAAAGCTCACTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCCACAGTATCTCTT 1047
|
|
|
QY 1026 ATTGCCAACATTGGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
|
|
|
Db 1048 ATTGCCAACATTGGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1107
|
|
|
QY 1086 GCTGCCAGTCCAAAGGAGGTGATCAATTG 1114
|
|
|
Db 1108 GCTGCCAGTCCAAAGGAGGTGATCAANTCG 1136
|
|
|
RESULT 6
AB017430 2097 bp mRNA linear PRI 06-MAR-1999
LOCUS Homo sapiens mRNA for kinesin-like DNA binding protein, complete cds.
ACCESSION AB017430 D38751
VERSION AB017430.2 GI:4519442
KEYWORDS Kid; kinesin-like DNA binding protein; kinesin family.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS 1 (sites)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tokai,N., Fujimoto-Nishiyama,A., Toyoshima,Y., Yonemura,S.,
Teukita,S., Inoue,J. and Yamamoto,T.
Kid, a novel kinesin-like DNA binding protein, is localized to
chromosomes and the mitotic spindle
EMBO J. 15 (3), 457-467 (1996)
PUBMED 96174806
REFERENCE 2 (bases 1 to 2097)
Tokai-Nishizumi,N. and Edamasu,M.
Direct Submission
Submitted (03-SEP-1998) Noriko Tokai-Nishizumi, The University of
Tokyo, The Institute of Medical Science, Department of Oncology;
4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
(E-mail:tokai@hgc.ims.u-tokyo.ac.jp, Tel:03-5449-5302,
Fax:03-5449-5413)
D38751:Submitted (02-Nov-1994)
Sequence updated (25-Feb-1999).
FEATURES Location/Qualifiers
```

source	1. .2097	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="pBskid"	
		/cell_line="MDA-MB453"	
		/cell_type="breast cancer"	
gene	1. .2097	/genes="kid"	
CDS	25. .2022	/gene="kid"	
		/codon_start=1	
		/product="kinesin-like DNA binding protein"	
		/protein_id="BA33019.2"	
		/db_xref="GI:4519443"	
		/translation="MAAGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPPARV	
		RVAVRLPFDVGTAGSDPPCVRGMDSCSLEIANWRNHQETLKYQDFGERSTQOD	
		IVAGVOPILRHLEGQNSVLAAGPTGAGKTHMLGSPPOGVI PALMDLLQLTRF	
		EGAEGRPWALSVTMSVLEIYQEKVLDLDPASGDVIREDCRGNILIPGLSKPISSF	
		ADFERHPLASNRRTVGATRLNORSRSHAVLLVKVDORERLAPFRQREKLYLIDLA	
		GSBNRRTGNKRLKESGAINSTSLFVLGKVDALNOGLPRVPRDSKUTRLQDLSLG	
		LSAHLILIANAPERFYDLSALNPAARSKVINRPFNLSLQPHALGPVKLSQKE	
		LLGPPPAKRAPPEEIEIGSPMAAPASQKSLPLQKLSMDPAMLERLLSLDRLL	
		ASQSGAPLLSTPKERVMVMTVEKLEIEIRLKTQKELEKAKMLAKAEKEKNC	
		PTWLRPLSHRTVGAELKKAIVMPLQIOEQASPNAEIHLKNGKRKRLKSLDAL	
		EPEKAEDEWELQISPELLAHGRQKILDLNLSGARDLRSLSQRIKPKKAQLIVGWEEL	
		HGPFQVEDLUEVEGITGKQWESFLKNIILGLAAGRCRGAS"	
ORIGIN			
Query Match	99.3%;	Score 1107.4;	DB 9; Length 2097;
Best Local Similarity	99.9%;	Pred. No. 2.6e-276;	
Matches 1108;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	6	GCOCGGCGCGCTCGACGACGACGAGCGACGCGAGATGCGCGAGCTTCAAGCGCGCG	65
DB	28	GCOCGGCGCGCTCGACGACGACGAGCGACGCGAGATGCGCGAGCTTCAAGCGCGCG	87
QY	66	ATCTCAGAGCTGCTCGCTCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCA	125
DB	88	ATCTCAGAGCTGCTCGCTCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCA	147
QY	126	GCTCGCTAAGGGTGGCTGTGCGACTCGCGGCATTTGTGGATGGAACAGCGGGAGCAAGT	185
DB	148	GCTCGCTAAGGGTGGCTGTGCGACTCGCGGCATTTGTGGATGGAACAGCGGGAGCAAGT	207
QY	186	GATCCCCCTGTGTGCGGGCATGGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAAC	245
DB	208	GATCCCCCTGTGTGCGGGCATGGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAAC	267
QY	246	CACGAGGAGCTCTCAATACCACTTTGATGCGCTTCTATGGGAGGAGGTACTTCAGCAG	305
DB	268	CACGAGGAGCTCTCAATACCACTTTGATGCGCTTCTATGGGAGGAGGTACTTCAGCAG	327
QY	306	GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCC	365
DB	328	GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGAGATGCC	387
QY	366	AGTGTGCTTGCTATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCCA	425
DB	388	AGTGTGCTTGCTATGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCCA	447
QY	426	GAGCAACTGGGGTGATCCCGCGGGCTCTCATGACCTCTCTGAGCTCACAAGGGAGGAG	485
DB	448	GAGCAACTGGGGTGATCCCGCGGGCTCTCATGACCTCTCTGAGCTCACAAGGGAGGAG	507
QY	486	GGTCCGAGGCGCGCCATGGCGCCCTTTGTGTACCAATGCTTTACCTAGAGATCTACCAG	545
DB	508	GGTCCGAGGCGCGCCATGGCGCCCTTTGTGTACCAATGCTTTACCTAGAGATCTACCAG	567
QY	546	GAGAGGATTTAGACCTTCCTGGACCTTCTCGGGAGACCTGGTAATCCGAGAGACTGC	605
DB	568	GAGAGGATTTAGACCTTCCTGGACCTTCTCGGGAGACCTGGTAATCCGAGAGACTGC	627
QY	606	CGGGGGAATATCTCTGATTCGGGTCTCTCCAGAACCCATAGTAGCTTTGCTGATTTT	665
DB	628	CGGGGGAATATCTCTGATTCGGGTCTCTCCAGAACCCATAGTAGCTTTGCTGATTTT	687
QY	666	GAGCGGCACTTCTCTGCGACGAGTGGAAATCGGACTGTAGGAGCCACCGGGCTCAACAG	725
DB	688	GAGCGGCACTTCTCTGCGACGAGTGGAAATCGGACTGTAGGAGCCACCGGGCTCAACAG	747
QY	726	CGCTCTCTCGGCACTGCTGCTGCTCTCTGCTCAAGGTGGACCGGAGACGTTTGGCC	785
DB	748	CGCTCTCTCGGCACTGCTGCTGCTCTCTGCTCAAGGTGGACCGGAGACGTTTGGCC	807
QY	786	CAATTTTCGCGACGAGGAGGAAATCTCTACCTGATTGACTTTGGCTGGGTTCAGAGGACAAC	845
DB	808	CAATTTTCGCGACGAGGAGGAGGAAATCTCTACCTGATTGACTTTGGCTGGGTTCAGAGGACAAC	867
QY	846	CGGCGCACAGGCAACAAAGGCGCTTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG	905
DB	868	CGGCGCACAGGCAACAAAGGCGCTTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG	927
QY	906	TTTGTCTCTGGCAAGTGTAGTGGCTGGAATCAGGGCTCCCTGCTGCTACCTTATCGG	965
DB	928	TTTGTCTCTGGCAAGTGTAGTGGCTGGAATCAGGGCTCCCTGCTGCTACCTTATCGG	987
QY	966	GACAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGGTGGGTTCAGCCACAGTATCCTT	1025
DB	988	GACAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGGTGGGTTCAGCCACAGTATCCTT	1047
QY	1026	ATTGCCAATTTGCCCTTCAGAGACGCTTCTACCTAGACAGATCTCCGCACTCAACTTT	1085
DB	1048	ATTGCCAATTTGCCCTTCAGAGACGCTTCTACCTAGACAGATCTCCGCACTCAACTTT	1107
QY	1086	GCTGCCAGTCCAGGAGGTGATCAATTG 1114	
DB	1108	GCTGCCAGTCCAGGAGGTGATCAATCG 1136	
RESULT 7			
BC028155		2117 bp	mRNA linear PRI 29-JUN-2004
LOCUS			
DEFINITION			Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:40049
			IMAGE:5241557), complete cds.
ACCESSION			BC028155
VERSION			BC028155.1 GI:20380446
KEYWORDS			MGC
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE			1 (bases 1 to 2117)
AUTHORS			Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
			Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
			Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
			Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
			Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
			Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
			Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S.,
			Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
			Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
			McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
			Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
			Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
			Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S.,
			Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
			Boffard G.C., Blakesley R.W., Touchman J.W., Green E.D.,
			Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
			Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E.,
			Schneerch A., Schein J.E., Jones S.J. and Marra M.A.
			Generation and initial analysis of more than 15,000 full-length
			human and mouse cDNA sequences
TITLE			
JOURNAL			Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED			12477932
REFERENCE			2 (bases 1 to 2117)

**AUTHORS**  
**TITLE**  
**JOURNAL**

Strausberg, R.  
Direct Submission  
Submitted (08-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

**REMARK**  
**COMMENT**

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Place: 62 Row: 1 Column: 3  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, Similarity but not identity to protein.

**FEATURES**  
**source**

1..2117  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:40049 IMAGE:5241557"  
/tissue\_type="Brain, fetal, whole pooled"  
/clone\_lib="NIH MGC\_121"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..2117  
/gene="KIF22"  
/note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"  
/db\_xref="LocusID:3835"  
/db\_xref="MIM:603213"  
6..2003  
/gene="KIF22"  
/codon\_start=1  
/product="kinesin family member 22"  
/protein\_id="AAH28155.1"  
/db\_xref="GI:20380447"  
/db\_xref="LocusID:3835"  
/db\_xref="MIM:603213"  
/translation="MAAGSTQORREMAAASAAISGACRCLSKIGATRRPPARV  
RVAVLRPEVDTAGASPPCVRGWDSCLLEIANRNHQTLYKQFDYFGERSTQD  
IYAGSVQVPLHLEQNASVLAQPTAGKTHMLGSPREPQVPIPRALLQLLTRE  
EGAEGRPALSVTMSLSEIYQEKVLDLDPASGDLVIREDCRGNLILPGLSQPISSF  
ADPERHFLPASNRVTGATRNQSSRSRAVLLVKVDQRELRAPRQREGKLYLIDLA  
GSEDNRTNKGRLKESGANTSLFVLGVVDALNQLPRVPVPSKLRLLQDSLG  
GSAHSILANTAPRRFVLDVTSALNFAARKEVINRPTNESLQPHALGPVLSQKE  
LIGPEAKRAGPEEEIGSPPEAPASAKSLPLKLSMDPAMLERLLSLDRLI  
ASQSQGAPLLSTPKRERMLKTVKNDLSIERLTKQKEBAKLAQAEKENHC  
ATMURPLSHRTVTGAKPLKAVVMPQLIQQAASPNABIIHKKNGKRLKESLDA  
EPKEAKEDVQLQSPPELLAHGRQKLDLNEGSARDLSLRIGPKKAQLIVGWREL  
HGPFQSVDELVERVITGKQMESFLKANILGLAAGRCAS"

**gene**

**CDS**

**ORIGIN**

Query Match 99.3%; Score 1107.4; DB 9; Length 2117;  
Best Local Similarity 99.9%; Pred. No. 2.6e-276;  
Matches 1108; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

6 GCCCGGGCGGCTCGACGACGAGGCGACGCGAGATGGCGGACGCTTTCAGCGGGCGG 65  
|  
|  
|  
9 GCCCGGGCGGCTCGACGACGAGGCGACGCGAGATGGCGGACGCTTTCAGCGGGCGG 68  
|  
|  
|  
66 ATCTCAGGAGCTGTGCTGGCTAAGCAAGATTGGAGCTACTCGTCTGCACCTCCA 125  
|  
|  
|  
69 ATCTCAGGAGCTGTGCTGGCTAAGCAAGATTGGAGCTACTCGTCTGCACCTCCA 128  
|  
|  
|  
126 GCTCGGTAAGGTCGCTGCGACGCTGGGCCATTTGTGGATGGAACACGCGGAGCAAGT 185  
|  
|  
|  
129 GCTCGTAAGGTCGCTGCGACGCTGGGCCATTTGTGGATGGAACACGCGGAGCAAGT 188  
|  
|  
|  
186 GATCCCCCTGTGTCGGGGCATGACAGCTCTCTCTAGAGATTGCTAACTGGAGGAAC 245  
|  
|  
|  
189 GATCCCCCTGTGTCGGGGCATGACAGCTCTCTCTAGAGATTGCTAACTGGAGGAAC 248  
|  
|  
|  
246 CACAGGAGACTCTCAAAATACAGTTTGTATGCTCTTATGGGAGAGGAGTACTACGACG 305  
|  
|  
|  
249 CACAGGAGACTCTCAAAATACAGTTTGTATGCTCTTATGGGAGAGGAGTACTACGACG 308  
|  
|  
|  
306 GACATCTATGAGGTTTCAGTCAGCCCATCTTAAGGCATCTTCTGAGGGGAGGAGTCC 365  
|  
|  
|  
309 GACATCTATGAGGTTTCAGTCAGCCCATCTTAAGGCATCTTCTGAGGGGAGGAGTCC 368  
|  
|  
|  
366 AGTGTGCTTGCCTATGAGCCCAACAGGAGCTGGGAAGAGCACAATGCTGGGCGACCCA 425  
|  
|  
|  
369 AGTGTGCTTGCCTATGAGCCCAACAGGAGCTGGGAAGAGCACAATGCTGGGCGACCCA 428  
|  
|  
|  
426 GAGCAACCTGGGTTGATCCCGGGCTCTCTATGGACCTCTCTGAGCTCAACAGGGAGGAG 485  
|  
|  
|  
429 GAGCAACCTGGGTTGATCCCGGGCTCTCTATGGACCTCTCTGAGCTCAACAGGGAGGAG 488  
|  
|  
|  
486 GGTGCGGAGGGCGGCGCATGGGCCCTTTCTGTCAACATGCTTACTAGAGATCTACACG 545  
|  
|  
|  
489 GGTGCGGAGGGCGGCGCATGGGCCCTTTCTGTCAACATGCTTACTAGAGATCTACACG 548  
|  
|  
|  
546 GAGAAGGTATTAGACCTCTCTGGACCTCTCTGGGAGACCTGGTAAATCCGAGAAGACTCC 605  
|  
|  
|  
549 GAGAAGGTATTAGACCTCTCTGGACCTCTCTGGGAGACCTGGTAAATCCGAGAAGACTCC 608  
|  
|  
|  
606 CGGGGAATATCTCTGATTCCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTT 665  
|  
|  
|  
609 CGGGGAATATCTCTGATTCCCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTT 668  
|  
|  
|  
666 GAGCGGCACTTCTGCGGAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACCG 725  
|  
|  
|  
669 GAGCGGCACTTCTGCGGAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACCG 728  
|  
|  
|  
726 CGCTCTCCCGCAGTCATGCTGTCTCTGCTCAAGGTGGACGCGGGAAGCTTTGGCC 785  
|  
|  
|  
729 CGCTCTCCCGCAGTCATGCTGTCTCTGCTCAAGGTGGACGCGGGAAGCTTTGGCC 788  
|  
|  
|  
786 CCAATTCGCCAGCAGAGGGGAAAACTCTACCTGATTGACTTTGGTGGGTGAGAGGACAA 845  
|  
|  
|  
789 CCAATTCGCCAGCAGAGGGGAAAACTCTACCTGATTGACTTTGGTGGGTGAGAGGACAA 848  
|  
|  
|  
846 CGGCGCACAGCAACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905  
|  
|  
|  
849 CGGCGCACAGCAACAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTG 908  
|  
|  
|  
906 TTTGCTCTGGGCAAGTGGTAGCTGCTGAATCAGGGCTCCCTCGTGTACTCTTATCGG 965  
|  
|  
|  
909 TTTGCTCTGGGCAAGTGGTAGCTGCTGAATCAGGGCTCCCTCGTGTACTCTTATCGG 968  
|  
|  
|  
966 GACAGCAAGCTCACTCGCTTATTCAGGACTCTCTCTGGTGGCTCAGGCCACAGTATCTCT 1025  
|  
|  
|  
969 GACAGCAAGCTCACTCGCTTATTCAGGACTCTCTCTGGTGGCTCAGGCCACAGTATCTCT 1028  
|  
|  
|  
1026 ATTGCGCAATTCGCTGAGAGAGCTTCTACCTAGACAGTCTCCGCACTCAACTTT 1085  
|  
|  
|  
1029 ATTGCGCAATTCGCTGAGAGAGCTTCTACCTAGACAGTCTCCGCACTCAACTTT 1088  
|  
|  
|  
1086 GCTGCCAGTCCAAAGGAGGTGATCAATTG 1114







ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1464)  
AUTHORS Beraud, C.  
TITLE Motor proteins and methods for their use  
JOURNAL Patent: US 6387644-A 7 14-MAY-2002;  
FEATURES Location/Qualifiers  
source 1..1464  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 92.9%; Score 1035.4; DB 6; Length 1464;  
Best Local Similarity 99.9%; Pred. No. 1.2e-257;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

78 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCCAGCTCCAGTTCGGTAAAG 137  
Db 4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCCAGCTCCAGTTCGGTAAAG 63  
Qy 138 GTGCTGTGGCACTGGCGGCAATTTGTGGATGGAACAGCGGGAGCAAGTGATCCCGCTGT 197  
Db 64 GTGCTGTGGCACTGGCGGCAATTTGTGGATGGAACAGCGGGAGCAAGTGATCCCGCTGT 123  
Qy 198 GTGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCAAGGAGACT 257  
Db 124 GTGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCAAGGAGACT 183  
Qy 258 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317  
Db 184 CTCAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243  
Qy 318 GTTCAGTGCAGCCATCTTAAGCAGCTGCTGGAAGGCGAGATGCGAGTGCTTGCC 377  
Db 244 GTTCAGTGCAGCCATCTTAAGCAGCTGCTGGAAGGCGAGATGCGAGTGCTTGCC 303  
Qy 378 TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCAGCAACCTGGG 437  
Db 304 TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCAGCAACCTGGG 363  
Qy 438 GTGATCCCGGGCTCTCATGGAACCTCTGAGTCTCAAGGGAGGAGGTCGCGAGGC 497  
Db 364 GTGATCCCGGGCTCTCATGGAACCTCTGAGTCTCAAGGGAGGAGGTCGCGAGGC 423  
Qy 498 CGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCGAGGAAGGTATTA 557  
Db 424 CGGCATCGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCGAGGAAGGTATTA 483  
Qy 558 GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGGAATATC 617  
Db 484 GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGGAATATC 543  
Qy 618 CTGATTCGGGTCTCTCCAGAACCCATCAGTACTGTTGCTGATTTTGAAGCGCACTTC 677  
Db 544 CTGATTCGGGTCTCTCCAGAACCCATCAGTACTGTTGCTGATTTTGAAGCGCACTTC 603  
Qy 678 CTGCGCAGCACTCGAATTCGAGCTCTAGGAGCCACCGGCTCAACCGAGCTCTCCCGC 737  
Db 604 CTGCGCAGCACTCGAATTCGAGCTCTAGGAGCCACCGGCTCAACCGAGCTCTCCCGC 663  
Qy 738 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACAGCGGGGAACGTTTGGCCCAATTTGCGCAG 797  
Db 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACAGCGGGGAACGTTTGGCCCAATTTGCGCAG 723  
Qy 798 CGAGAGGAAATCTTACTGATTCAGTTCGCTGGGTGAGAGGCAACCGCGGCACAGGC 857  
Db 724 CGAGAGGAAATCTTACTGATTCAGTTCGCTGGGTGAGAGGCAACCGCGGCACAGGC 783  
Qy 858 AACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCTGGC 917  
Db 784 AACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCTGGC 843  
Qy 918 AAGTGGTAGATGGGCTGAATCAGGGCCTCCCTCGTGTACCTTATTCGGGACAGCAAGCTC 977

844 AAGTGGTAGATGGGCTGAATCAGGGCCTCCCTCGTGTACCTTATTCGGGACAGCAAGCTC 903  
Qy 978 ACTCGCCTATTGAGGACTCTCTGGGTGGTTCAGCCACAGTATCTTATTCGCAACATT 1037  
Db 904 ACTCGCCTATTGAGGACTCTCTGGGTGGTTCAGCCACAGTATCTTATTCGCAACATT 963  
Qy 1038 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCGAGTCC 1097  
Db 964 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCGAGTCC 1023  
Qy 1098 AAGGAGGTGATCAATTG 1114  
Db 1024 AAGGAGGTGATCAATCG 1040

BC003427 2086 bp mRNA linear ROD 29-JUN-2004  
IMAGE:2615715), complete cds.  
BC003427  
VERSION BC003427.1 GI:13097359  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2086)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.B.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
human and mouse cDNA sequences  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 2086)  
Strausberg, R.  
Direct Submission  
Submitted (20-FEB-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-x@mail.nih.gov](mailto:cgabs-x@mail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLML)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amc@bcm.tmc.edu](mailto:amc@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.







```
Db 65130 TCTGACTTCGAGAACGACTTCTCTTCTAGTCAGTAGAAATCGAGTCTTAGGAGCTACTCTG 65071
QY 717 CTCAACAGCGGCTCTCCCGCAGTCATGCTGCTCTCTGCTCAAGTCAGGACCGGGAA 776
Db 65070 CTTAACAGCGGCTCTGCGGTAGTCATGACGCGCTCTTGTCAAGGTAGATCAGCATGAA 65011
QY 777 CGTTTGGCCCAATTCGCCAGCGAGAGGAGAACTCTACTGATTCGACTTGGCTGGGTCA 836
Db 65010 CGTTTGACTCCATTTCCGCGAGTGGGAAGGAAACTCTACTTATTGCTTGGCTGGTTCA 64951
QY 837 GAGGACAAC-----GCGCACAGGCAACAGGCGCTTCGGCTAAAGAGAGATGGAGCCAT 892
Db 64950 GAGGACAACCGTCAGTCGACAGGCAACAGGCGCATTCCTGCTCAAGAGAGATGGAGCCAT 64891
QY 893 CAACACTCCCTGTTGTTGCTCTGCGCAAGTGGTGTAGATGCGCTGAATCAGGCGCTCCCTCG 952
Db 64890 CAACACTCCCTGTTGTTGCTCTGCGCAAGTGGTGTAGATGCGCTGAATCAGGCGCTCCCTCA 64831
QY 953 TGTACCTTATCGGAGCAGCAAGCTCACTCGCCTATTGTCAGGACTCTCTGGGTGGCTCAGC 1012
Db 64830 TATACCATACGAGCAGCAGGCTCACTCGTCTGTTGCAAGACTCTCTGAGAGGCTCAGC 64771
QY 1013 CCACAGTATCCTTATTCGCAACATTCGCCCTGAGACGCTTCTACCTAGACACAGTCTC 1072
Db 64770 TCATAGCATCCTCATTTGTCAGCATTTGCTCTGAAAGATGTTTTTACCTGGATACATCTC 64711
QY 1073 CGCACTCAACTTTGCTGCGCAGGTCGCAAGGAGGTGATCAA 1111
Db 64710 AGCATTAACCTTCACTGCTAGGTCGCAAGGAGGTGATTA 64672
```

## RESULT 13

```
AC101752/c
LOCUS AC101752
DEFINITION Mus musculus chromosome 1, clone RP24-344C18, complete sequence.
ACCESSION AC101752.14 GI:46518680
VERSION HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 164759)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 1, clone RP24-344C18
Unpublished
```

```
2 (bases 1 to 164759)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cook,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosett,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
```

REFERENCE  
AUTHORS

```
3 (bases 1 to 164759)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepe,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivier,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
```

REFERENCE  
AUTHORS

```
4 (bases 1 to 164759)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArelano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Olivier,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
```

## COMMENT

On Apr 23, 2004 this sequence version replaced gi:45504330.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submission@broad.mit.edu](mailto:sequence_submission@broad.mit.edu)

----- Project Information

Center project name: L17344

Center clone name: 344\_C\_18

FEATURES  
source

Location/Qualifiers  
1..164759  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clone="RP24-344C18"  
/clone\_lib="RPMI-24 Male Mouse BAC"  
repeat\_region 1..1655  
/rpt\_family="L1\_MM"

```

1. .4
/note="clone boundary
clone_end:SF6
site:MboI"
1767..1787
/rpt_family="AT_rich"
2115..2157
/rpt_family="AT_rich"
2236..2258
/rpt_family="AT_rich"
3252..3420
/rpt_family="AT_rich"
3464..3588
/rpt_family="Lx7"
/rpt_family="Lx7"
3605..3666
/rpt_family="CT-rich"
3673..3731
/rpt_family="(TTC)n"
3732..3911
/rpt_family="(TTC)n"
4000..4180
/rpt_family="CT-rich"
4202..4363
/rpt_family="Lx7"
4529..4765
/rpt_family="MTB"
7234..7442
/rpt_family="B3"
8445..8509
/rpt_family="T-rich"
complement(8512..8591)
/rpt_family="MLT11"
9336..9632
/rpt_family="MLT13"
complement(9800..11396)
/rpt_family="L1_MM"
complement(11397..17527)
/rpt_family="L1F"
complement(18235..18777)
/rpt_family="Lx"
18775..18890
/rpt_family="Lx"
19260..19949
/rpt_family="RMER19A"
20118..20297
/rpt_family="MLT1D"
complement(20386..20639)
/rpt_family="Lx2"
20634..20839
/rpt_family="Lx4"
20876..21793
/rpt_family="RLTRI13D"
21799..21907
/rpt_family="RMER13A"
21918..22433
/rpt_family="Lx4"
22591..22651
/rpt_family="(TG)n"
22733..24486
/rpt_family="L1_MM"
24494..24850
/rpt_family="L1_MM"
27139..27166
/rpt_family="AT_rich"
27737..27834
/rpt_family="MLT1H"
28328..28506
/rpt_family="MLT1K"
28513..28541
/rpt_family="AT_rich"
28813..28858
/rpt_family="AT_rich"
29627..30186

```

```

repeat_region
/rpt_family="L1_MM"
30465..30578
/rpt_family="T-rich"
complement(30581..34457)
/rpt_family="L1_MM"
complement(34455..35895)
/rpt_family="L1"
complement(36134..36736)
/rpt_family="Lx4"
36749..36920
/rpt_family="Lx2"
37617..37751
/rpt_family="(TA)n"
37958..37986
/rpt_family="(GAA)n"

Query Match 61.1%; Score 681.4; DB 10; Length 164759;
Best Local Similarity 79.7%; Pred. No. 8.9e-166;
Matches 844; Conservative 0; Mismatches 206; Indels 9; Gaps 3;

Qy 57 GCGCGCGGATCTCAGAGCTGCTCGCTGCGGCTAAGCAAGATTGGAGCTACTCGTCGT 116
Db 146137 GCCTCAGCCACTCAGGACCGGGTCTTTGT---ATAAGCAAGGTAGGATTTGGCTCGT 146081
Qy 117 CCACCTCCAGTCGCGTAAGGTGCTGCGACTGCGGCCATTGTGGATTGAACAGCG 176
Db 146080 CTGCTTCTAGTGAGTAAGGTGGCTATATATTGTGCCGTGTTTGGATTGAAGAGACA 146021
Qy 177 GGAGCAAGTGATCCCCCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAAC 236
Db 146020 GAAGCGAAGGAGCACCCCTGTGTACGAGCCATACAGAGCTGCTCTCTGAAGTGCTAAC 145961
Qy 237 TGGAGGAACCAACAGGAGACTCTCAATACAGTTTGATGCTTCTATGGGAGAGAGT 296
Db 145960 TGGAGGAATAACAGGAGACTATCAAAATATCAGTTTGATGCTTCTTCCATGGAGAGAGC 145901
Qy 297 ACTCAGCAGGACATCTATCAGGTTTCTAGTGAGGCTTCTTCAAGGCACTTCTCGAAGGG 356
Db 145900 ACTCAGTGGACATCTATGTTAGTTTATGACGCCATCTAAGCACTTCTTGGAGAGG 145841
Qy 357 CAGAAATCCAGTGTGCTTGTGCTATGGACCCACAGGAGCTGGGAGAGCGACAAATGCTG 416
Db 145840 CAGAAATCCAGTGTGCTTGTGCTATGGACCCAGTGGGCGAGGAGAGA--CACACAATGTTG 145783
Qy 417 GGCAGCCCGAGAGCACTGGGGTGATCCCGGGGCTCTATGGACCTCTCGACCTCACA 476
Db 145782 GGCAGCCCGAGAGCACTGGAGTGTGCTTCTCAGACTCTCCGACCTCTCGACCTAACA 145723
Qy 477 AGGAGGAGGCTGCCGAGGGCGGCCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAG 536
Db 145722 AGGAGAGAGATGCTGAGGGCTGGCCATGGGACATTTCTGTGTTTGTGCTATTTAGAT 145663
Qy 537 ATCTACAGGAGAGGATTTAGACCTCTCGACCTCTTGGAGACCTTGGGAGACCTGGTAAATCCGA 596
Db 145662 CTCTAACAGGAAATAGTATTAGACCTCTTGGATCTGCAATCAGAAAGACCTAGTGTCCGA 145603
Qy 597 GAAGACTGCGGGGGAATATCTTGATTCGGGTCTCTCCGAGGAGCCATCAGTAGCTTT 656
Db 145602 GAAGGCTGACGAGGGGAAATCTTGATTCGAGGCCCTCACAGAAAGCCCATACATAGTTTC 145543
Qy 657 GCTGATTTGAGCGGCACTTCTGCGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGG 716
Db 145542 TCTGACTTCGAGAGCACTTCTTCTTAGTCAGTAGAATCGATTTGTAGGAGCTACTCTG 145483
Qy 717 CTCAACAGCGCTCTCCCGAGTCATGCTGTGCTCTCTGTTCAAGGTGAGCCAGCGGAA 776
Db 145482 CTTAACAGCGCTCTTCCCGTAGTCATGCGCGCTCTTGTGTTCAAGGTAGATCAGCATGAA 145423
Qy 777 GGTGCGCCCATTTCCGCGAGGAGGAGGAAACTCTACCTGATTGACTTGGCTGGGTCA 836
Db 145422 CGTTTGACTCTCATTTCCCGAGGAGGAAACTCTACCTTATTTGATTTGGCTGGGTCA 145363
Qy 837 GAGGACAAAC---GGCGCACAGGCAACAAAGGGGCTTTCCGCTAAAGAGAGTGGAGCCAT 892

```

**TITLE**  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 297639)  
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collumore, A., Cook, A., Cooke, P., Corum, B., DeArillano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-pierre, N., Hafez, N., Haggopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Katam, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,

<hr/>							
/clone_lib="RPL1-24 Male Mouse BAC"							
<hr/>							
ORIGIN							
Query Match		58.7%;		Score 654.2;		DB 2; Length 297639;	
Best Local Similarity		78.1%;		Pred. No. 1e-158;			
Matches 827; Conservative		0;		Mismatches 223;		Indels 9; Gaps 3;	
<hr/>							
QY	57	GCGGGGCGCATCTCAGGAGCTGGTGCCTTCGGCTTAAGCAAGATTGGAGCTACTGTCTCGT					116
Db	1205	GCGTCAGCCACCCTCAGGACCGGCTTTGCT---ATAAGCAAGTAGGATGGCTGTCTCGT					1149
<hr/>							
QY	117	CACCTTCAGCTCGGCTAAGGTGGCTGFCGCACTCGGCCCAATTTGTGAATGAAACAGCG					176
Db	1148	CTGCTTCTAGCTTGGAATGAAGGTGGCTATPATAATTGTGCTGTCTTTGTGAAGAAGACA					1089

Qy	177	GGAGCAAGTGATCCCCCTGTGTGGGGCATGACAGCTGCTCTAGAGATTGCTAAC	236
Ds	1088	GAAGCGAGGAGCACCCTGTGTACGACCATGACAGCTGCTCTTGAAGTGGCTAAC	1029
Qy	237	TGGAGGAACCAACGAGGAGACTCTCAATACAGTTTGTATGCTCTATGGGAGAGAGT	296
Ds	1028	TGGAGGAATACCAAGGAGACTATCAATATCAGTTTGTATGCTCTTCCATGGAGAGAGC	969
Qy	297	ACTCAGCAGGACATCTATGAGGTTTCAGTGAGGCCCATCTTAAGCACTTCTCTGAAGGG	356
Ds	968	ACTCAGTGGGACATCTATGAGGTTTGTAGTACAGCCCATACTAAGAACTTCTCTGAAGGG	909
Qy	357	CAGAAATGCACTGCTTGTCCCTATGAGCCACAGGAGCTGGGAAGAGCACAATGCTG	416
Ds	908	CAGAAATGCACTGCTTGTCCCTATGAGCCACAGGAGCTGGGAAGAGCACAATGCTG	851
Qy	417	GGCAGCCAGAGCAACTCTGGGGTATCCCGGGCTCTCATGAGACCTCTCTGACCTCACA	476
Ds	850	GGCAGCCAGAGCAACTCTGGGGTATCCCGGGCTCTCATGAGACCTCTCTGACCTCACA	791
Qy	477	AGGAGGAGGGTGGGAGGGCGGCGCATGGGCGCTTCTGTGACCATGTCTTACCTAGAG	536
Ds	790	AGGAGGAGAAATCTGAGGGCTGGCCATGAGACATTTCTGTGTTGTCTATTTAGAT	731
Qy	537	ATCTACAGGAGAAAGTATTAGACTCTCTGACCTCTCTCGGAGACCTGGTAATCCGA	596
Ds	730	CTCTAACAGGAAATGATTATGACCTCTTGATCTCTCTATAGTGTACCTAATGTCGA	671
Qy	597	GAAGACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGAGCCCATCAGTAGCTTT	656
Ds	670	CGGCGAGGACGAGGAGAAATCCTGATTCAGGCGCTCACAAGAGCCCATCAGTAGTTT	611
Qy	657	GCTGATTTGAGCGGCACTCTCTCCAGCCAGTCGAAATCGGACTGTAGGACCCACCGG	716
Ds	610	TCTGACTTCGAGAGGCACTCTCTCTAGTCACTAGTAAATCGAGTTGTAGGAGTACTCTG	551
Qy	717	CTCAACAGCGCTCTCTCCGAGTCATGCTGTGCTCTCTGTCGAGGTGGACCGCGGAA	776
Ds	550	CTTAACAGCGCTCTTGGCGTAGTCATGCGCGCTCTTGGTCAGAGTAGATCAGCATGAA	491
Qy	777	CGTTTGGCCCATTTGCGCAGGAGAGGAAAACTTACCTGATTTGCTGGCTGGGTCA	836
Ds	490	CGTTTGAATCTTCTGCGCAGTGGGAGGAAAACTTACCTTATTTGATTTGGCTGGTCA	431
Qy	837	GAGCAGACCC---GGGCGCAGGCAACAGGGCTTGGCTAAAGAGAGTGGAGCAT	892
Ds	430	GAGCAGACCCGTCACTGCGCAGGCAACAGGGCAATTTCTGCTCAAGAGAGTGGAGCAT	371
Qy	893	CAACACCTCTCTGTTTGTCTGGGCAAGTGGTATGCGCTGAATCAGGGCTCTCCCTCG	952
Ds	370	CAACACCTCTCTTGTGTTGCTGGGCAAGTGGTATGCGCTGAATCAGGGCTCTCCCTCA	311
Qy	953	TGTACCTTATGGGACAGCAAGCTCATCTGCGCTATTCAGGAGCTCTCTGGGTGGCTCAGC	1012
Ds	310	TATACCATACAGGACAGCAGGCTCACTCGTCTGTTGCAAGACTCTCTGAGGCTCAGC	251
Qy	1013	CCACAGTATCTTATTTGCCAATGTCCTCCCTGACAGAGCTTCTTACCTAGACAGTCTC	1072
Ds	250	TCATAGATCTCTCATTTGTCAGCATTTGCTCTGAAAGATGTTTTTACCTGGATACATCTC	191
Qy	1073	CGCACTCAACTTTGCTGTCAGGTCGAAGGAGGTGATCAA	1111
Ds	190	AGCATTAACCTTCACTGCTAGTCCAGGAGGTGATTTAA	152

RESULT 15  
 AC102127  
 LOCUS AC102127  
 DEFINITION Mus musculus chromosome 1, clone RP23-202A19, complete sequence.  
 ACCESSION AC102127  
 VERSION AC102127.11  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 196674)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 1, clone RP23-202A19  
 Unpublished

REFERENCE  
AUTHORS

2 (bases 1 to 196674)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,  
 Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Collings,M., Collins,S., Collamore,A., Cook,A.,  
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Goyette,M., Graham,L., Grand-pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 McLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,  
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollard,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Vainour,J., Vo,A., Wilson,B., Wu,X., Wyman,D., Yel,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission

TITLE  
JOURNAL

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 196674)

REFERENCE  
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,S., Atachchi,H.M., Barna,N., Bastien,V.,  
 Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J.,  
 Choepel,Y., Collamore,A., Cook,A., Cooke,P., Corum,B.,  
 DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
 Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D.,  
 Galagan,J., Gardyna,S., Graham,L., Grand-pierre,N., Hafez,N.,  
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
 MacLean,C., MacDonald,P., Major,J., Manning,J., Matthews,C.,  
 McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
 Zimmer,A. and Zody,M.  
 Direct Submission

TITLE  
JOURNAL

Submitted (10-APR-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 196674)

REFERENCE  
AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,S., Atachchi,H.M., Barna,N., Bastien,V.,  
 Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J.,  
 Choepel,Y., Collamore,A., Cook,A., Cooke,P., Corum,B.,  
 DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
 Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D.,  
 Galagan,J., Gardyna,S., Graham,L., Grand-pierre,N., Hafez,N.,  
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
 MacLean,C., MacDonald,P., Major,J., Manning,J., Matthews,C.,  
 McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
 Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
 Zimmer,A. and Zody,M.  
 Direct Submission



Qy 439 TGATCCCGGGCTCTCATGACCTCTGAGCTCAAGGAGGAGGTTGCCGAGGCC 498  
 Db 101932 TGATTCCTTAGCCCTCTCATAGACCTCTGCAGATAACAAGGGAAGAAAGTGCTGAAGGCT 101991  
 Qy 499 GGCCATGGGCGCTTCTGTGTCACCATGCTTACCTAGAGATCTACCGAGGAAGGTATTAG 558  
 Db 101992 GGCCATGGGACATTTCTGTAGGATGTCCTATTAGAGGTTCTACCGAGGAAGGTATTAG 102051  
 Qy 559 ACCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGGGGAATATCC 618  
 Db 102052 ACCTCTTATATCTCTGCATCAGGAGACCTAGTGATATGAGAAGATTGTGAGGGAACATCC 102111  
 Qy 619 TGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGGGCACTTCC 678  
 Db 102112 TGATTCAGTTCTCACCAGGACCCCAATCTGCTTCTCTGACTTTGAGCAGTACTT-C 102170  
 Qy 679 TGCCAGGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACGCGCTCTCTCCGCA 738  
 Db 102171 TGCCAGCCTGTAGAAATCAAGTTGTAGAGACTACTCTGCTTAAGCAGTGTCTTACCATA 102230  
 Qy 739 GTCATGTGTGCTCTCTGGTCAAGGTGACACAGCGGGAAAGTTGGCCCCCATTTGCCGAGC 798  
 Db 102231 -TCATTCAATGCTCTTTATCAAGGTGGATCAGCATGGACATTTGACTCTGTTTCACTAGT 102289  
 Qy 799 GAGAGGGAAGAACTCTACTGATTCGACTTTGGTGGTTCAGAGGCAACCGGGGCACAGGCA 858  
 Db 102290 GAAAGGGAAGAACTCTACCATATTGATTGGTTGGCTCAGAGGCAACCAACA-AGGA 102348  
 Qy 859 ACAAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCA 918  
 Db 102349 ACCAGTGCATTCAACTCAAGAGAGTGGCGCCATCAACACT-GCTTCTTGTAAGTGGCA 102407  
 Qy 919 AAGTGGTAGATGCGCTGAATCAGGGCTCTCCCTGCTGTACCTTATCGGGACAGCAAGCTCA 978  
 Db 102408 AAGTGGAGTTGCATCATATTAGGGCTCCCTCTCTATATTATCTGGGACAGAAAAATCA 102467  
 Qy 979 CTCGCTATTGCAGGACTCTCTGGTGGCTCAGGCCACAGTATCTTTATTGCCAACATTG 1038  
 Db 102468 TTTCCTATTGCCAGACTCTCTGGAAGCTCAGCTCATACATCATATTACCACACTG 102527  
 Qy 1039 CCCCTGAGAGAGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCCA 1098  
 Db 102528 CTCCTGAAAGACATTTTACTTTGGAATCATTTCTGAGCATTTGAAGTCACTGCTAGGTCCA 102587  
 Qy 1099 AGGAGTGATCAATTG 1114  
 Db 102588 AGGAGCTGATTAATTG 102603

Search completed: November 10, 2004, 13:17:54  
 Job time : 4934.36 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 13:18:18 ; Search time 781.387 Seconds  
(without alignments)  
10087.639 Million cell updates/sec

Title: US-10-797-893-7  
Perfect score: 1464  
Sequence: 1 atgggtcgctgtcggttaag.....accattgtccacaatgtga 1464

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464	100.0	1464	17	US-10-797-893-7
2	1461	99.8	1538	17	US-10-797-893-3
3	1447	98.8	2099	16	US-10-334-143-102
4	1443.2	98.6	2104	9	US-09-925-300-288
5	1038.4	70.9	1041	17	US-10-797-893-5
6	1035.4	70.7	1115	17	US-10-797-893-1
7	480	32.8	491	9	US-09-960-253-131
8	376.4	25.7	386	10	US-09-803-719-1937
9	365	24.9	24525	13	US-10-087-192-1180
10	265	18.1	464	10	US-09-918-995-31705
11	262.4	17.9	24291	13	US-10-087-192-1177
12	233.2	15.9	531	15	US-10-029-386-10603

13	230.4	15.7	232	15	US-10-029-386-24303	Sequence 24303, A
14	185.6	12.7	2095	16	US-10-108-260A-249	Sequence 249, App
15	156.8	10.7	3374	16	US-10-287-226-301	Sequence 301, App
16	138.6	9.5	4108	9	US-09-883-096-1	Sequence 1, Appli
17	137.2	9.4	2675	16	US-10-287-226-305	Sequence 305, App
18	136.6	9.3	1152	9	US-09-883-096-6	Sequence 6, Appli
c 19	134	9.2	2607	17	US-10-437-963-1460	Sequence 1460, Ap
c 20	127.6	8.7	482	16	US-10-152-319A-498	Sequence 498, App
21	123.2	8.4	2698	16	US-10-425-114-25129	Sequence 25129, A
22	121.8	8.3	3661	16	US-10-311-642-1	Sequence 1, Appli
23	121.6	8.3	3034	18	US-10-425-115-148256	Sequence 148256,
24	121.4	8.3	1014	9	US-09-883-096-3	Sequence 3, Appli
25	113.6	7.8	987	16	US-10-311-642-3	Sequence 313, App
26	112.2	7.7	4292	16	US-10-287-226-313	Sequence 720, App
27	110.2	7.5	740	14	US-10-076-555-720	Sequence 384, App
28	107	7.3	259	14	US-10-076-555-384	Sequence 828, App
29	107	7.3	657	14	US-10-076-555-828	Sequence 56548, A
30	106.2	7.3	1998	17	US-10-437-963-56548	Sequence 135546,
31	101	6.9	2264	18	US-10-425-115-135546	Sequence 586, App
32	101	6.9	3258	16	US-10-302-172-586	Sequence 23, Appli
33	100.2	6.8	3657	16	US-10-336-472-23	Sequence 57, Appli
34	100.2	6.8	3657	16	US-10-236-417-57	Sequence 1087, Ap
35	98.2	6.7	4724	15	US-10-172-118-1087	Sequence 1087, Ap
36	98.2	6.7	4724	16	US-10-342-887-1087	Sequence 201, App
c 37	94.4	6.4	2217	15	US-10-104-047-201	Sequence 22135, A
38	91.4	6.2	3390	16	US-10-425-114-22135	Sequence 2359, Ap
39	91.4	6.2	3407	18	US-10-739-930-2359	Sequence 46568, A
40	91.4	6.2	3433	18	US-10-425-115-46568	Sequence 34, Appli
41	90.4	6.2	4913	15	US-10-133-937-34	Sequence 34, Appli
42	90.4	6.2	4913	16	US-10-159-563-34	Sequence 27001, A
43	88.8	6.1	1295	15	US-10-369-493-27001	Sequence 762, App
44	88.4	6.0	3159	15	US-10-104-047-762	Sequence 1053, Ap
45	86.2	5.9	6927	15	US-10-172-118-1053	

## ALIGNMENTS

RESULT 1  
US-10-797-893-7  
; Sequence 7, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1464  
; TYPE: DNA  
; ORGANISM: Human  
US-10-797-893-7

Query Match 100.0%; Score 1464; DB 17; Length 1464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGGTCGCTGTGCGGCTAACAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGGTA 60  
Db 1 ATGGGTCGCTGTGCGGCTAACAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGGTA 60  
Qy 61 AGGGTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGAGCAAGTATGATCCCCC 120  
Db 61 AGGGTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGAGCAAGTATGATCCCCC 120

QY 121 TGTGTGGGGGATGAGACAGTCTCTCTAGAGATTGCTAACTGGAGGAAACACAGGAG 180  
Db 121 TGTGTGGGGGATGAGACAGTCTCTCTAGAGATTGCTAACTGGAGGAAACACAGGAG 180  
QY 181 ACTCTCAATACACAGTTTCATGCTCTATATGGGAGAGAGTACTCAGCAGACATCTAT 240  
Db 181 ACTCTCAATACACAGTTTCATGCTCTATATGGGAGAGAGTACTCAGCAGACATCTAT 240  
QY 241 GCAGGTTTCAGTGCAGGCCCATCTTAAAGGCACTTGTCTGAAAGGCGAGAAATCCAGTGTGTT 300  
Db 241 GCAGGTTTCAGTGCAGGCCCATCTTAAAGGCACTTGTCTGAAAGGCGAGAAATCCAGTGTGTT 300  
QY 301 GCTATGGAACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGCGAGCCCGAGCAACT 360  
Db 301 GCTATGGAACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGCGAGCCCGAGCAACT 360  
QY 361 GGGGTGATCCCGGGGCTCTCATGAGCCTCTCTGAGCTCACAAGGAGGAGGCTCCGAG 420  
Db 361 GGGGTGATCCCGGGGCTCTCATGAGCCTCTCTGAGCTCACAAGGAGGAGGCTCCGAG 420  
QY 421 GCGCGGCCATGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTTACAGGAAAGTA 480  
Db 421 GCGCGGCCATGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTTACAGGAAAGTA 480  
QY 481 TTAGACTCTCTGACCTCTTCCGAGACCTGGTAATCCGAGAACTGCGCGGGGAAT 540  
Db 481 TTAGACTCTCTGACCTCTTCCGAGACCTGGTAATCCGAGAACTGCGCGGGGAAT 540  
QY 541 ATCTCTGATCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600  
Db 541 ATCTCTGATCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600  
QY 601 TTCTCTGACGCCAGTCCGAATCGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCC 660  
Db 601 TTCTCTGACGCCAGTCCGAATCGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCC 660  
QY 661 CGCAGTCATGCTGTCTCTGCTCAAGTGGACGAGGGAAGTTTGGCCCATTTCCG 720  
Db 661 CGCAGTCATGCTGTCTCTGCTCAAGTGGACGAGGGAAGTTTGGCCCATTTCCG 720  
QY 721 CAGCAGAGGGAATCTTACCTGATTTGCTGGCTGGGTGAGAGCAACCGCGGCACA 780  
Db 721 CAGCAGAGGGAATCTTACCTGATTTGCTGGCTGGGTGAGAGCAACCGCGGCACA 780  
QY 781 GGCACAAAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTG 840  
Db 781 GGCACAAAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTG 840  
QY 841 GGCACAAAGTGTAGATGCGCTGAATCAGGGCTCTCTGCTGATCTTATCGGACAGCAAG 900  
Db 841 GGCACAAAGTGTAGATGCGCTGAATCAGGGCTCTCTGCTGATCTTATCGGACAGCAAG 900  
QY 901 CTCACCTCGCTATTGACAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCAAC 960  
Db 901 CTCACCTCGCTATTGACAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATGCAAC 960  
QY 961 ATTGCCCCGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTGTCTGAGG 1020  
Db 961 ATTGCCCCGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTGTCTGAGG 1020  
QY 1021 TCCAAAGGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCCTCATGCTTGGGA 1080  
Db 1021 TCCAAAGGAGTGTATCAATCGGCTTTTACCAATGAGAGCTGAGCCTCATGCTTGGGA 1080  
QY 1081 CCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140  
Db 1081 CCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140  
QY 1141 CTTGAGGAAGAGGATTTGGAGCCCTGAGCCCATGAGCTCCAGCTCTGCTCCAG 1200  
Db 1141 CTTGAGGAAGAGGATTTGGAGCCCTGAGCCCATGAGCTCCAGCTCTGCTCCAG 1200  
QY 1201 AAACCTCAGCCCCCTACAGAGCTAAGCAGCATGGACCGGCCATGCTGGAGCGCTCCTC 1260

Db 1201 AAACCTCAGCCCCCTACAGAGCTAAGCAGCATGGACCGGCCATGCTGGAGCGCTCCTC 1260  
QY 1261 AGCTTTGGACCGTCTGCTTTCCTCCAGGGGAGCCAGGGGGCCCTCTCTTTGAGTACCCCA 1320  
Db 1261 AGCTTTGGACCGTCTGCTTTCCTCCAGGGGAGCCAGGGGGCCCTCTCTTTGAGTACCCCA 1320  
QY 1321 AAGCGAGAGCGATGGTGTCTAATGAGACAGTACAGAGAGGAGGAGCTAGAGATTGAGAGG 1380  
Db 1321 AAGCGAGAGCGATGGTGTCTAATGAGACAGTACAGAGAGGAGGAGCTAGAGATTGAGAGG 1380  
QY 1381 CTTTAAAGCAAGCAAAAGAACTGGAGGCCAAGATTTGGCCCAAGAGGCTGAGGAAAG 1440  
Db 1381 CTTTAAAGCAAGCAAAAGAACTGGAGGCCAAGATTTGGCCCAAGAGGCTGAGGAAAG 1440  
QY 1441 GAGAACCAATTGTCCCAATGTGA 1464  
Db 1441 GAGAACCAATTGTCCCAATGTGA 1464

RESULT 2  
US-10-797-893-3  
; Sequence 3, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1538  
; TYPE: DNA  
; ORGANISM: Human  
US-10-797-893-3

Query Match 99.8%; Score 1461; DB 17; Length 1538;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGTCTGCTGCTGCTAAGCAAGATTGGAGCTACTCTGCTGCTCACTCCAGCTCGCGTAAGG 63  
Db 78 GGTCTGCTGCTGCTAAGCAAGATTGGAGCTACTCTGCTGCTCACTCCAGCTCGCGTAAGG 137  
QY 64 GTGGCTGTGCGACTGGGCCATTTGATGGAACAGCGGGAGCAAGTATCCCCCTCT 123  
Db 138 GTGGCTGTGCGACTGGGCCATTTGATGGAACAGCGGGAGCAAGTATCCCCCTCT 197  
QY 124 GTGGGGGCGATGAGACAGTCTCTCTAGAGATTGCTTAACTGGAGAAACACAGAGAGCT 183  
Db 198 GTGGGGGCGATGAGACAGTCTCTCTAGAGATTGCTTAACTGGAGAAACACAGAGAGCT 257  
QY 184 CTCAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTATGCA 243  
Db 258 CTCAATACCAAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTATGCA 317  
QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTTGCTGGAAGGCGAGATGCCAGTGTGCTTGC 303  
Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTTGCTGGAAGGCGAGATGCCAGTGTGCTTGC 377  
QY 304 TATGAGCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGAGCAAGCAACTGGG 363  
Db 378 TATGAGCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGAGCAAGCAACTGGG 437  
QY 364 GTGATCCCGGGGCTCTATGAGCTCTGAGCTCTCAAGGGAGGAGGTTCCGAGGGC 423

Db 438 GTGATCCCGGGCTCTCATGGAAGCTCTGTGAGCTCAAGAGGAGGGGTGCCAGGGC 497  
Qy 424 CGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACCAAGAGAAAGTATTA 483  
Db 498 CGGCCATGGGCCCTTTCTGTCAACATGCTTACCTAGAGATCTACCAAGAGAAAGTATTA 557  
Qy 484 GACCTCTGGAACCTTCTGTTCGGGAGACCTGTGTAATCCGAGAAGACTGCGGGGGAATATC 543  
Db 558 GACCTCTGGAACCTTCTGTTCGGGAGACCTGTGTAATCCGAGAAGACTGCGGGGGAATATC 617  
Qy 544 CTGATTCGGGGTCTCTCCAGAACCCATCAGTAGCTTTCGTAATTTGAGCGGCACTTC 603  
Db 618 CTGATTCGGGGTCTCTCCAGAACCCATCAGTAGCTTTCGTAATTTGAGCGGCACTTC 677  
Qy 604 CTGCAGCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGAGCGCTCTCCCGC 663  
Db 678 CTGCAGCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGAGCGCTCTCCCGC 737  
Qy 664 AGTCATGCTGTCTCTGCTCAAGGTGAGACAGCGGGAAGCTTTGGCCCCCATTTTCGCCAG 723  
Db 738 AGTCATGCTGTCTCTGCTCAAGGTGAGACAGCGGGAAGCTTTGGCCCCCATTTTCGCCAG 797  
Qy 724 CGAGAGGAAACTCTACCTGATTTGACTTGGCTGGGTCAAGAGCAACCCGCGCACAGGC 783  
Db 798 CGAGAGGAAACTCTACCTGATTTGACTTGGCTGGGTCAAGAGCAACCCGCGCACAGGC 857  
Qy 784 AACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGC 843  
Db 858 AACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGC 917  
Qy 844 AAAGTGTAGATGCGCTGAAATCAGGGCTCTCCCTGCTGTACCTTATCGGGACAGCAAGCTC 903  
Db 918 AAAGTGTAGATGCGCTGAAATCAGGGCTCTCCCTGCTGTACCTTATCGGGACAGCAAGCTC 977  
Qy 904 ACTCGCCTATTCAGGACTCTCTGGGTGGCTCAGCCACAGTAGTATTCCTTATGCGCAAT 963  
Db 978 ACTCGCCTATTCAGGACTCTCTGGGTGGCTCAGCCACAGTAGTATTCCTTATGCGCAAT 1037  
Qy 964 GCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGTCC 1023  
Db 1038 GCCCTCAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGTCC 1097  
Qy 1024 AAGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGACGCTCATGCTTTGGGACCT 1083  
Db 1098 AAGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGACGCTCATGCTTTGGGACCT 1157  
Qy 1084 GTTAAGCTGTCTAGAAAGATTCCTTGTGTCACAGAGCAAGAGAGCCCGAGGCGCT 1143  
Db 1158 GTTAAGCTGTCTAGAAAGATTCCTTGTGTCACAGAGCAAGAGAGCCCGAGGCGCT 1217  
Qy 1144 GAGGAGAGGAGATTGGAGCCCTGAGCCATGSCAGCTCCAGCTCTGCTCCAGAAA 1203  
Db 1218 GAGGAGAGGAGATTGGAGCCCTGAGCCATGSCAGCTCCAGCTCTGCTCCAGAAA 1277  
Qy 1204 CTCAGCCCCCTACAGAGCTAAGCAGCATGGACCCCGGCCATGTGGAGCGCTCTCCAGC 1263  
Db 1278 CTCAGCCCCCTACAGAGCTAAGCAGCATGGACCCCGGCCATGTGGAGCGCTCTCCAGC 1337  
Qy 1264 TTGAGCCGTCTGTCTCCAGAGGAGCCAGGGGCGCCCTCTGTGTAGTACCCCAAG 1323  
Db 1338 TTGAGCCGTCTGTCTCCAGAGGAGCCAGGGGCGCCCTCTGTGTAGTACCCCAAG 1397  
Qy 1324 CGAGAGCGGATGCTTAATGAGACAGTAGAGAGAGAGGACCTTAGAGATTGAGAGCTT 1383  
Db 1398 CGAGAGCGGATGCTTAATGAGACAGTAGAGAGAGAGGACCTTAGAGATTGAGAGCTT 1457  
Qy 1384 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCCAGAGGCTGAGGAAAGGAG 1443  
Db 1458 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCCAGAGGCTGAGGAAAGGAG 1517  
Qy 1444 AACCATTTGCCCAATGTGA 1464  
Db 1518 AACCATTTGCCCAATGTGA 1538

## RESULT 3

US-10-334-143-102  
; Sequence 102, Application US/10334143  
; Publication No. US2004009549A1  
; GENERAL INFORMATION:  
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
; APPLICANT: SUDARSANAM, SUCHA  
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
; FILE REFERENCE: 038602/1543  
; CURRENT APPLICATION NUMBER: US/10/334,143  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: 60/343,169  
; PRIOR FILING DATE: 2001-12-31  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 102  
; LENGTH: 2099  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-334-143-102

Query Match 98.8%; Score 1447; DB 16; Length 2099;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1458; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 4 GGTGCTGTTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGGTAAGG 63  
Db 101 GGTGCTGTTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGGTAAGG 160  
Qy 64 GTGCTGTGCGACTCGGCCATTTGTGGATGGAAACAGCGGGAGCAAGTATCCCCCTGT 123  
Db 161 GTGCTGTGCGACTCGGCCATTTGTGGATGGAAACAGCGGGAGCAAGTATCCCCCTGT 220  
Qy 124 GTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCAACAGAGACT 183  
Db 221 GTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAAACCAACAGAGACT 280  
Qy 184 CTCAATATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243  
Db 281 CTCAATATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 340  
Qy 244 GGTTCAGTGCAGCCCATCTTAAGGCATCTGCTGGAGGGGAGATGCGAGTGTGCTTGGC 303  
Db 341 GGTTCAGTGCAGCCCATCTTAAGGCATCTGCTGGAGGGGAGATGCGAGTGTGCTTGGC 400  
Qy 304 TATGGACCCACAGGAGCTGGGAAAGACGACACAATGCTGGGAGCCCGAGAGCAACCTGGG 363  
Db 401 TATGGACCCACAGGAGCTGGGAAAGACGACACAATGCTGGGAGCCCGAGAGCAACCTGGG 460  
Qy 364 GTGATCCCGGGCTCTCTATGGACCTCTCTGACCTCAAGAGGAGGAGGTGCGGAGGC 423  
Db 461 GTGATCCCGGGCTCTCTATGGACCTCTCTGACCTCAAGAGGAGGAGGTGCGGAGGC 520  
Qy 424 CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGAGAGAGGTATTA 483  
Db 521 CGGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGAGAGAGGTATTA 580  
Qy 484 GACCTCTGAGACCTCTCTCGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAATATC 543  
Db 581 GACCTCTGAGACCTCTCTCGGAGACCTGGTAATCCGAGAAGACTGCCGGGGAATATC 640  
Qy 544 CTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 603  
Db 641 CTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 700  
Qy 604 CTGCCAGCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGAGCGCTCTCCCGC 663  
Db 701 CTGCCAGCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGAGCGCTCTCCCGC 760  
Qy 664 AGTCATGCTGTGCTCTCTGCTCAAGGTGGACCGGGGAAGCTTTGGCCCCCATTTTCGCCAG 723

Db 761 ACTCATGCTGCTCTCTGCTCAAGTGGACAGCGGGAAGCTTTGGCCCCATTTGGCCAG 820  
QY 724 CGAGAGGGAACCTCTACCTGATGATGCTTGGGTGAGAGCAACCGCGGCGACAGGC 783  
Db 821 CGAGAGGGAACCTCTACCTGATGATGCTTGGGTGAGAGCAACCGCGGCGACAGGC 880  
QY 784 AACAGGGCTTCGGCTAAAGAGAGTGAGGCATCAACACCTCCCTGTTGCTCTGGC 843  
Db 881 AACAGGGCTTCGGCTAAAGAGAGTGAGGCATCAACACCTCCCTGTTGCTCTGGC 940  
QY 844 AAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGACACAGCTC 903  
Db 941 AAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGACACAGCTC 1000  
QY 904 ACTGCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAACATT 963  
Db 1001 ACTGCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAACATT 1060  
QY 964 GCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGGTCC 1023  
Db 1061 GCCCTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGGTCC 1120  
QY 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGCCTCATGCTTTGGACCT 1083  
Db 1121 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCAGCCTCATGCTTTGGACCT 1180  
QY 1084 GTTAAGCTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGCCCT 1143  
Db 1181 GTTAAGCTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGCCCT 1240  
QY 1144 GAGGAGAGAGATTGGGAGCCCTGAGCCATGAGAGCTTCAGCCTTCGCTCCAGAAA 1203  
Db 1241 GAGGAGAGAGATTGGGAGCCCTGAGCCATGAGAGCTTCAGCCTTCGCTCCAGAAA 1300  
QY 1204 CTCAGGCCCTCAGAAAGCTTAAGCAGCATGAGAGCTTCAGCCTTCGCTCCAGC 1263  
Db 1301 CTCAGGCCCTCAGAAAGCTTAAGCAGCATGAGAGCTTCAGCCTTCGCTCCAGC 1360  
QY 1264 TTGAGCCGTCTGCTGCTCCAGAGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323  
Db 1361 TTGAGCCGTCTGCTGCTCCAGAGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1420  
QY 1324 CGAGAGCGATGCTGCTAATGAGACAGTAGAGAGAGAGAGCTAG-AGATTGAGAGGCT 1382  
Db 1421 CGAGAGCGATGCTGCTAATGAGACAGTAGAGAGAGAGAGCTAGTAGATTGAGAGGCT 1480  
QY 1383 TAAGACGAAAGCAAAAGAACTGGAGGCCAAGATGTTGGGCCAGAGGCTGAGGAAAGGA 1442  
Db 1481 TAAGACGAAAGCAAAAGAACTGGAGGCCAAGATGTTGGGCCAGAGGCTGAGGAAAGGA 1540  
QY 1443 GAACCATTTGCCCAATG 1461  
Db 1541 GAACCATTTGCCCAATG 1559

RESULT 4  
US-09-925-300-288  
; Sequence 288, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 288  
; LENGTH: 2104  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (44)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (497)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (1323)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-300-288  
Query Match 98.6%; Score 1443.2; DB 9; Length 2104;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1450; Conservative 7; Mismatches 1; Indels 1; Gaps 1;  
QY 4 GGTCCCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTAAGG 63  
Db 67 GGTCCCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTAAGG 126  
QY 64 GTGGCTGTGCGACTCGGCGCAATTTGTGGATGGAAACAGCGGGAGCAAGTGATCCCCCTGT 123  
Db 127 GTGGCTGTGCGACTCGGCGCAATTTGTGGATGGAAACAGCGGGAGCAAGTGATCCCCCTGT 186  
QY 124 GTGCGGGCATGAGCAGCTGCTCTTAGAGATTGTAACTGGAGGAAACCAACAGAGACT 183  
Db 187 GTGCGGGCATGAGCAGCTGCTCTTAGAGATTGTAACTGGAGGAAACCAACAGAGACT 246  
QY 184 CTCAAATACCAAGTTTGTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243  
Db 247 CTCAAATACCAAGTTTGTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 306  
QY 244 GTTCAGTGCAGCCCATCTAAGGCATTTGCTGGAAGGCGCAATGCCAGTGTGCTTGC 303  
Db 307 GTTCAGTGCAGCCCATCTAAGGCATTTGCTGGAAGGCGCAATGCCAGTGTGCTTGC 366  
QY 304 TATGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGCAGAGCAACTGGG 363  
Db 367 TATGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGCAGAGCAACTGGG 426  
QY 364 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGAGAGAGGTGCCGAGGSC 423  
Db 427 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGAGAGAGGTGCCGAGGSC 486  
QY 424 CGGCCATGGGCTTTTCTGTCAACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483  
Db 487 CGGCCATGGGCTTTTCTGTCAACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 546  
QY 484 GACCTCTTGAGACCTTCTGCTGGGAGACTGTTAAATCCGAGAGAGACTGCCGGGGGAATATC 543  
Db 547 GACCTCTTGAGACCTTCTGCTGGGAGACTGTTAAATCCGAGAGAGACTGCCGGGGGAATATC 606  
QY 544 CTGATTCGGGTCTCTCCAGAGCCCATCAGTACTGCTTGTGATTTTTCAGCGGCACTTC 603  
Db 607 CTGATTCGGGTCTCTCCAGAGCCCATCAGTACTGCTTGTGATTTTTCAGCGGCACTTC 666  
QY 604 CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCACCCCGCTCAACCCAGCGCTCTCCCGC 663  
Db 667 CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCACCCCGCTCAACCCAGCGCTCTCCCGC 726  
QY 664 AGTCATGCTGCTCTGCTGATGAGTGGACGAGGGAACGTTTGGCCCCCATTTGCCAG 723  
Db 727 AGTCATGCTGCTCTGCTGATGAGTGGACGAGGGAACGTTTGGCCCCCATTTGCCAG 786  
QY 724 CGAGGAGGAAACTCTTACCTGATTTGCTTGGCTGAGAGCAACACCGGCGCAGGC 783  
Db 787 CGAGGAGGAAACTCTTACCTGATTTGCTTGGCTGAGAGCAACACCGGCGCAGGC 846  
QY 784 AACAGGGCTTCGGCTAAAGAGAGTGAGGCATCAACACCTCCCTGTTGCTCTGGGC 843



```

RESULT 6
US-10-797-893-1
; Sequence 1, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-1

Query Match      70.7%; Score 1035.4; DB 17; Length 1115;
Best Local Similarity 99.9%; Pred. No. 8e-311;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4  GGTCCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCCGCTAAGG 63
DB      78  GGTCCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCCGCTAAGG 137

QY      64  GTCCGCTGTGCGACTGCGGCCATTGTGGATGGAACACGCGGAGCAAGTATCCCCCTGT 123
DB     138  GTCCGCTGTGCGACTGCGGCCATTGTGGATGGAACACGCGGAGCAAGTATCCCCCTGT 197

QY     124  GTCCGGGGCATGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGGACCAACAGGAGACT 183
DB     198  GTCCGGGGCATGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGGACCAACAGGAGACT 257

QY     184  CTCAAATACCACTTTGATGCTCTTATGGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 243
DB     258  CTCAAATACCACTTTGATGCTCTTATGGGGAGAGGAGTACTCAGCAGGAGCATCTATGCA 317

QY     244  GGTTCAGTGAGCCCATCTTAAGGCACTTGTCTGGAGGGGAGAGTCCAGTGTGCTTGCC 303
DB     318  GGTTCAGTGAGCCCATCTTAAGGCACTTGTCTGGAGGGGAGAGTCCAGTGTGCTTGCC 377

QY     304  TATGGACCCACAGAGCTGGGAGAGCGCACAAATGCTGGGAGCCCGAGCAACCTGGG 363
DB     378  TATGGACCCACAGAGCTGGGAGAGCGCACAAATGCTGGGAGCCCGAGCAACCTGGG 437

QY     364  GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAGGAGGAGGGGTGCGAGGGC 423
DB     438  GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAGGAGGAGGGGTGCGAGGGC 497

QY     424  CGGCCATGGGCCCTTTCTGTACCATGTCTTAAGTCTTAGAGATCTACAGGAGAGGATTA 483
DB     498  CGGCCATGGGCCCTTTCTGTACCATGTCTTAAGTCTTAGAGATCTACAGGAGAGGATTA 557

QY     484  GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCGGGGGGAATATC 543
DB     558  GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCGGGGGGAATATC 617

QY     544  CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTTGGAGCGGCATTC 603
DB     618  CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTGCTGATTTTGGAGCGGCATTC 677

QY     604  CTGCCAGCAGTGGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGCGCTCTCTCCCGC 663
DB     678  CTGCCAGCAGTGGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGCGCTCTCTCCCGC 737

```

```

QY     664  AGTCATGCTGTGCTCCTCTCAAGGTGGACACGCGGAAAGCTTTGGSCCCCATTTCCGCCAG 723
DB      738  AGTCATGCTGTGCTCCTCTCAAGGTGGACACGCGGAAAGCTTTGGSCCCCATTTCCGCCAG 797

QY     724  CGAGAGGGGAAAACTCTTACCTGATTGACTTTGGCTGGGTCTAGAGGACAAACGCGGCACAGGC 783
DB      798  CGAGAGGGGAAAACTCTTACCTGATTGACTTTGGCTGGGTCTAGAGGACAAACGCGGCACAGGC 857

QY     784  AACAAGGGCCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCTCTGTTGTTCTCTGGGC 843
DB     858  AACAAGGGCCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCTCTGTTGTTCTCTGGGC 917

QY     844  AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903
DB     918  AAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 977

QY     904  ACTCGCTTATTGACGACTCTCTTGGGTGGCTCAGCCCAAGTATCTTATTCGAACATT 963
DB     978  ACTCGCTTATTGACGACTCTCTTGGGTGGCTCAGCCCAAGTATCTTATTCGAACATT 1037

QY     964  GCGCCTGAGAGAGCGTTTACCTAGACACAGTCTCGGACCTCAACTTTGCTGCGAGGTCC 1023
DB    1038  GCGCCTGAGAGAGCGTTTACCTAGACACAGTCTCGGACCTCAACTTTGCTGCGAGGTCC 1097

QY    1024  AAGGAGGTGATCAATCG 1040
DB    1098  AAGGAGGTGATCAATCG 1114

RESULT 7
US-09-960-253-131
; Sequence 131, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.556
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 131
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-131

Query Match      32.8%; Score 480; DB 9; Length 491;
Best Local Similarity 99.8%; Pred. No. 2.3e-138;
Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY     936  AGCCCAACAGTATCTTATTTGCCAAACATTCCTGAGAGACGCTTCTACTAGACACAGT 995
DB      936  AGCCCAACAGTATCTTATTTGCCAAACATTCCTGAGAGACGCTTCTACTAGACACAGT 60

QY     996  CTCGCACTCAACTTTGCTGCCAGGTCCAAGGAGGTGATCAATCGGCCCTTTTACCAATGA 1055
DB      996  CTCGCACTCAACTTTGCTGCCAGGTCCAAGGAGGTGATCAATCGGCCCTTTT-CCAATGA 119

QY    1056  GAGCCTGAGCCTCATGCTTGGACCTTAAAGCTGCTCAGAAAGAAATGCTTGGTCC 1115
DB    1200  GAGCCTGAGCCTCATGCTTGGACCTTAAAGCTGCTCAGAAAGAAATGCTTGGTCC 179

QY    1116  ACCAGAGGCAAGAGAGAGCCCGGAGGCTCAGGAGAGGAGATTTGGAGCCCTTGGAGCCAT 1175
DB    1800  ACCAGAGGCAAGAGAGAGCCCGGAGGCTCAGGAGAGGAGATTTGGAGCCCTTGGAGCCAT 239

QY    1176  GGAGCTCCAGGCTCTGCTTCCAGAACTCAGACCCCTTACAGAGCTAAGAGCATGGA 1235
DB    2400  GGAGCTCCAGGCTCTGCTTCCAGAACTCAGACCCCTTACAGAGCTAAGAGCATGGA 299

```

Qy 1236 CCGGCGCATGTGGAGCGCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGGAGCCA 1295  
 Db 300 CCGGCGCATGTGGAGCGCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGGAGCCA 359  
 Qy 1296 GGGGGCCCTCTGTGTAGTACCCCAAGAGGAGCGGATGTCTAATGAAGACAGTAGA 1355  
 Db 360 GGGGGCCCTCTGTGTAGTACCCCAAGAGGAGCGGATGTCTAATGAAGACAGTAGA 419  
 Qy 1356 AGAAGAGACCTAGAGATTGAGAGGCTTAAGACGAAGCAAAAGAACTGGAGGCGCAAGAT 1415  
 Db 420 AGAAGAGACCTAGAGATTGAGAGGCTTAAGACGAAGCAAAAGAACTGGAGGCGCAAGAT 479  
 Qy 1416 GTTGGCCCGAGAA 1427  
 Db 480 GTTGGCCCGAGAA 491

RESULT 8

US-09-803-719-1937  
 ; Sequence 1937, Application US/09803719  
 ; Publication No. US20030044783A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Williams, Lewis T.  
 ; APPLICANT: Escobedo, Jaime  
 ; APPLICANT: Innis, Michael A.  
 ; APPLICANT: Garcia, Pablo Dominguez  
 ; APPLICANT: Sudduth-Klinger, Julie  
 ; APPLICANT: Reinhard, Christoph  
 ; APPLICANT: Giese, Klaus  
 ; APPLICANT: Randazzo, Filippo  
 ; APPLICANT: Kennedy, Giulia C.  
 ; APPLICANT: Pot, David  
 ; APPLICANT: Kassam, Altaf  
 ; APPLICANT: Lamson, George  
 ; APPLICANT: Drmanac, Radoje  
 ; APPLICANT: Crkvenjakov, Radomir  
 ; APPLICANT: Dickson, Mark  
 ; APPLICANT: Drmanac, Snezana  
 ; APPLICANT: Labat, Ivan  
 ; APPLICANT: Leshkowitz, Dena  
 ; APPLICANT: Kita, David  
 ; APPLICANT: Garcia, Veronica  
 ; APPLICANT: Jones, Lee William  
 ; APPLICANT: Stache-Crain, Birgit  
 ; TITLE OF INVENTION: Human Genes and Gene Products  
 ; FILE REFERENCE: 1624.002  
 ; CURRENT APPLICATION NUMBER: US/09/803,719  
 ; CURRENT FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 60/188,609  
 ; PRIOR FILING DATE: 2000-03-09  
 ; NUMBER OF SEQ ID NOS: 2396  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1937  
 ; LENGTH: 386  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-803-719-1937

Query Match 25.7%; Score 376.4; DB 10; Length 386;  
 Best Local Similarity 98.4%; Pred. No. 3.3e-106;  
 Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1073 CTTTGGGACTGTGAAGCTGTCTCAGAAAGAAATTTGTTGTCACACAGGCGCAAGAGAG 1132  
 Db 1 CGTTTGTGTCGTTTAAAGCTGTCTCAGAAAGAAATTTGTTGTCACACAGGCGCAAGAGAG 60  
 Qy 1133 CCGAGGCCCTAGAGAGAGAGATTGGGAGCCCTGAGGCCCATGGCAGCTCCAGCCCTCTG 1192  
 Db 61 CCGAGGCCCTAGAGAGAGAGATTGGGAGCCCTGAGGCCCATGGCAGCTCCAGCCCTCTG 120  
 Qy 1193 CTTCCAGAAACTCAGCCCCCTACAGAAAGCTTAAGCAGCATGGACCCCGCCATGCTGGAGC 1252

Db 121 CTTCCAGAAACTCAGCCCCCTACAGAAAGCTTAAGCAGCATGGACCCCGCCATGCTGGAGC 180  
 Qy 1253 GCCTCCTCAGCTTGGACCGCTGCTGCTCCAGGGGAGCCAGGGGCCCTCTCTGTGA 1312  
 Db 181 GCCTCCTCAGCTTGGACCGCTGCTGCTCCAGGGGAGCCAGGGGCCCTCTCTGTGA 240  
 Qy 1313 GTACCCCAAGCGAGAGCGGATGTCTAATGAAGACAGTAAAGAGAGAGACCTTAGAGA 1372  
 Db 241 GTACCCCAAGCGAGAGCGGATGTCTAATGAAGACAGTAAAGAGAGAGACCTTAGAGA 300  
 Qy 1373 TTGAGAGGCTTAAGACCAAGCAAAAGAACTGGAGGCCCAAGATGTTGGCCCAAGAGGCTG 1432  
 Db 301 TTGAGAGGCTTAAGACCAAGCAAAAGAACTGGAGGCCCAAGATGTTGGCCCAAGAGGCTG 360  
 Qy 1433 AGGAAAGGAGAACCATTTGTCCCA 1458  
 Db 361 AGGAAAGGAGAACCATTTGTCCCA 386

RESULT 9

US-10-087-192-1180  
 ; Sequence 1180, Application US/10087192  
 ; Publication No. US20020182586A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morris, David W.  
 ; APPLICANT: Engelhard, Eric K.  
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
 ; TITLE OF INVENTION: CANCER  
 ; FILE REFERENCE: 529452000122  
 ; CURRENT APPLICATION NUMBER: US/10/087,192  
 ; CURRENT FILING DATE: 2002-03-01  
 ; PRIOR APPLICATION NUMBER: US 09/747,377  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: US 09/798,586  
 ; PRIOR FILING DATE: 2001-03-02  
 ; NUMBER OF SEQ ID NOS: 2059  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1180  
 ; LENGTH: 24525  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-087-192-1180

Query Match 24.9%; Score 365; DB 13; Length 24525;  
 Best Local Similarity 73.8%; Pred. No. 5.1e-102;  
 Matches 597; Conservative 0; Mismatches 0; Indels 212; Gaps 2;  
 Qy 476 AGGTATTAGACCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGG 535  
 Db 2313 AGGTATTAGACCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGG 2372  
 Qy 536 GGAATATCTGATTCGGGCTCTCCAGAGAGCCCATCAGTAGCTTGTGATTTGAGC 595  
 Db 2373 GGAATATCTGATTCGGGCTCTCTCCAGAGAGCCCATCAGTAGCTTGTGATTTGAGC 2432  
 Qy 596 GGCACCTTCCTGCCAGCCAGTCGAAATCGGACTGTAGGAGCACCCGCTCAACACAGCGCT 655  
 Db 2433 GGCACCTTCCTGCCAGCCAGTCGAAATCGGACTGTAGGAGCACCCGCTCAACACAGCGCT 2492  
 Qy 656 CTTCCCGCAGTCATGCTGCTCTCTGTC-----AGGTGACCGAG 685  
 Db 2493 CTTCCCGCAGTCATGCTGCTCTCTGTCAGAGTGGAGCCGACAGAGGGCGAGGAGCCT 2552  
 Qy 686 -----AGGTGACCGAG 696  
 Db 2553 GGAAGAGCCAGGAGCCTAGCTAAGCAGAGACCTTTGTTCTTACCCCGGAGTGGACAG 2612  
 Qy 697 CGGGAACGTTTGGCCCCCATTTTCGCGAGAGAGGAAACTCTACCTGATTGACTTGCT 756  
 Db 2613 CGGGAACGTTTGGCCCCCATTTTCGCGAGAGAGGAAACTCTACCTGATTGACTTGCT 2672  
 Qy 757 GGGTCAAGAGGACCAACCGCGCACAGGCAACAAGGGCCTTCGGCTAAAGAGAGTGGAGCC 816



Db 2673 GGGTCAGAGCAACACGGCGGCACAGGCAACAGGGCCCTTCGGCTAAAAGAGGTGAGCC 2732  
Qy 817 ATCAACACCTCCCTGTTTCTCTGGGCAAAAGTGGTAGATGCGCTGAATCAGGGCCTCCCT 876  
Db 2733 ATCAACACCTCCCTGTTTCTCTGGGCAAAAGTGGTAGATGCGCTGAATCAGGGCCTCCCT 2792  
Qy 877 CGTGACCTTATCGGACAGCAAGCTCACTCGCCTATTG----- 915  
Db 2793 CGTGACCTTATCGGACAGCAAGCTCACTCGCCTATTG----- 2852  
Qy 916 ----- 915  
Db 2853 AGGGAAGAGGGGCTGCAGAGAGGTTCTCAGGCTGCTGTGGGTGGGATAGCAGT 2912  
Qy 916 -----CAGGACTC 923  
Db 2913 TGAGGCATAGGAAGGTGGGCTTCTGACCCACCCACTGCTGCTCACCCTCAGGACTC 2972  
Qy 924 TCTGGGTGGCTAGCCACAGTATCTTATTCGCAACATTGCCCTTGAGAGACGCTTCTA 983  
Db 2973 TCTGGGTGGCTAGCCACAGTATCTTATTCGCAACATTGCCCTTGAGAGACGCTTCTA 3032  
Qy 984 CTTAGACAGCTCTCGGCACTCAACTTGTGCGAGTCCAGGAGTGAATCGGCC 1043  
Db 3033 CTTAGACAGCTCTCGGCACTCAACTTGTGCGAGTCCAGGAGTGAATCGGCC 3092  
Qy 1044 TTTTACCAATGAGAGCTGCGCTCATG 1072  
Db 3093 TTTTACCAATGAGAGCTGCGCTCATG 3121

RESULT 10  
US-09-918-995-31705  
; Sequence 31705, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 31705  
; LENGTH: 464  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(464)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-31705

Query Match 18.1%; Score 265; DB 10; Length 464;  
Best Local Similarity 77.6%; Pred. No. 1.5e-71; Indels 1; Gaps 1;  
Matches 333; Conservative 0; Mismatches 95  
Qy 126 GCGGGGATGACAGCTGCTCTCTAGAGATTGCTTAATCGGAGGAACACCCAGGAGACTCT 185  
Db 37 GCCTGGAAATTCAAATATCGGCTCGATAGAGTGAACCGGCCGAACCAACCGGATAATCT 96  
Qy 186 CAAATACAGTTTATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGACG 245  
Db 97 CCCTTACCTGTGCTTTCCTTTTGCAGCCATAGGAGGACTGCGCTGACATGTACTTCC 156  
Qy 246 TTCAGTGACGCCATCTTAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTCTTGCTA 305  
Db 157 TTGACCGTGGACCGGTGCGCACTGCGCACTCTCTTCGTAATTAACGAGGAGCTGGCTA 216  
Qy 306 TGGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGCAGCAACCTGGGGT 365

Db 217 CGGTGCTGAGGAGTGCATTAC-ATCTCGATGCTTGGCAGCCAGAGCAACCTGGGGT 275  
Qy 366 GATCCCGGGGCTCTCATGACCTCTCGAGCTTCAAGAGGAGAGGTCGCCGAGGGCCG 425  
Db 276 GATCCCGGGGCTCTCATGACCTCTCGAGCTTCAAGAGGAGAGGTCGCCGAGGGCCG 335  
Qy 426 GCCATGGGCTTCTCTGACCATGCTTACCTAGAGATCTTACAGGAGAGGTATTAGA 485  
Db 336 GCCATGGGCTTCTCTGACCATGCTTACCTAGAGATCTTACAGGAGAGGTATTAGA 395  
Qy 486 CCTCTGTCACCTCTCTCGGAGACCTGGTAATCCGAGAGACTGCGGGGGAATATCCT 545  
Db 396 CCTCTGTCACCTCTCTCGGAGACCTTGGTAATCCGAGAGACTGCGGGGGAATATCCT 455  
Qy 546 GATTCCGGG 554  
Db 456 GATTCCGGG 464

RESULT 11  
US-10-087-192-1177  
; Sequence 1177, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1177  
; LENGTH: 24291  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(24291)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-1177

Query Match 17.9%; Score 262.4; DB 13; Length 24291;  
Best Local Similarity 73.8%; Pred. No. 3.9e-70; Indels 76; Gaps 1;  
Matches 385; Conservative 0; Mismatches 61  
Qy 476 AGGTATTAGACCTCTCGGAGACCTGTTGAGAGACTGTTGATTCGAGAGACTGCGGG 535  
Db 2939 AGGTATTAGACCTCTCGGAGACCTGTTGAGAGACTGTTGATTCGAGAGACTGCGGG 2998  
Qy 536 GGAATATCTGATTCGGGTCTCTCCGAGAGCCATCAGTAGCTTGTGATTTTGGC 595  
Db 2999 GGAACATCTGATTCGAGGCTTCAACAGAGCCATCAGGAGCTTCTGACTTCGAGC 3058  
Qy 596 GGCACCTCTGCGGAGCTGGAATCGGAGCTGAGGAGCCACCGGCTCAACAGGCGCT 655  
Db 3059 AGCACCTCTTCCAGCCAGTGAATTCGAGCTGTAGGAGCCACCGGCTTAACAGGCT 3118  
Qy 656 CTTCCCGCAGTCTGCTGCTCTGCTCTGCTCA----- 685  
Db 3119 CTTCCCGTGTAGTACGAGTGTCTTGTTCAGGTAAGGCCCGCTGACAGAAAGGCTGGG 3178  
Qy 686 -----AGTGTACAGGCGG 699  
Db 3179 AAGCCCTGAAATGTGAGTGTGGGTCTCTTCTCTTCCCTAGGTAGTACGAGCT 3238  
Qy 700 GAACGTTTGGCCCACTTCGCCAGCGAGAGGGGAAACTCTACCTGATTGCTGGCTGGG 759



Db 3239 GAACTTTGACTCCATTTCGCCAGCGGGAAGAAACTCTACCTTATTGATTGGCTGGT 3298  
Qy 760 TCAGAGACAACCGCGCACAGGCAACAAGGCGCTTCGGCTAAAGAGAGTGGAGCCATC 819  
Db 3299 TCAGAGACAACCGTCGCACAGGCAACAGGCGCATTCGGCTCAAGAGAGCGGAGCCATC 3358  
Qy 820 AACACCTCCCTGTTTGTCTCGGCAAAAGTGTAGATCGCTGAATCAGGCGCTCCCTCGT 879  
Db 3359 AACACCTCCCTGTTTGTACTGGGCAAAAGTGTGTGATGATTAACACGAGGCGCTCCCTCGC 3418  
Qy 880 GTACCTTATCGGACAGCAAGCTCACTCGCTTATTCAGGAC 921  
Db 3419 ATACCATACCGGACAGCAAGCTCACTCGCTGTCGAGGTC 3460

## RESULT 12

US-10-029-386-10603  
; Sequence 10603, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AROMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 10603  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7  
; OTHER INFORMATION: NT HIT: AB017334.1, EVALUATE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUATE 2.00e-36  
; OTHER INFORMATION: EST\_HUMAN HIT: BF975048.1, EVALUATE 1.00e-126  
US-10-029-386-10603

Query Match 15.9%; Score 233.2; DB 15; Length 531;  
Best Local Similarity 98.7%; Pred. No. 1.2e-61;  
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 684 CAAGGTGGACCGAGCGGAACGTTTGGCCCCCATTTCCGCGAGGAGGGAAGAACTCTACCT 743  
Db 34 CCAGGTGGACCGAGCGGAACGTTTGGCCCCCATTTCCGCGAGGAGGGAAGAACTCTACCT 93  
Qy 744 GATTGACTTGGCTGGGTTCAGAGGACAACCGCGCGCACAGGCAACAAGGCGCTTCGGCTAAA 803  
Db 94 GATTGACTTGGCTGGGTTCAGAGGACAACCGCGCGCACAGGCAACAAGGCGCTTCGGCTAAA 153  
Qy 804 AGAGAGTGGAGGCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATGGCTGAA 863  
Db 154 AGAGAGTGGAGGCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATGGCTGAA 213  
Qy 864 TCAGGGGCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTTATTCAGGAC 921  
Db 214 TCAGGGGCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTTATTCAGGAC 271

## RESULT 13

US-10-029-386-24303  
; Sequence 24303, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AROMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 24303  
; LENGTH: 232  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7  
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUATE 5.00e-37  
; OTHER INFORMATION: EST\_HUMAN HIT: BE382882.1, EVALUATE 1.00e-125  
; OTHER INFORMATION: NT HIT: gi6453817, EVALUATE 1.00e-125  
US-10-029-386-24303

Query Match 15.7%; Score 230.4; DB 15; Length 232;  
Best Local Similarity 99.6%; Pred. No. 6.4e-61;  
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 688 GTGACCAAGCGGGAACGTTTGGCCCCCATTTCCGCGAGGAGGGAAGAACTCTACCTGATT 747  
Db 1 GTGACCAAGCGGGAACGTTTGGCCCCCATTTCCGCGAGGAGGGAAGAACTCTACCTGATT 60  
Qy 748 GACTTGGCTGGGTTCAGAGGACAACCGCGCGCACAGGCAACAAGGCGCTTCGGCTAAAAGAG 807  
Db 61 GACTTGGCTGGGTTCAGAGGACAACCGCGCGCACAGGCAACAAGGCGCTTCGGCTAAAAGAG 120  
Qy 808 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATGGCTGAATCAG 867  
Db 121 AGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATGGCTGAATCAG 180  
Qy 868 GGCCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTTATTCGAGG 919  
Db 181 GGCCTCCCTCGTGTACTTATCGGACAGCAAGCTCACTCGCTTATTCGAGG 232

## RESULT 14

US-10-108-260A-249  
; Sequence 249, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 249  
; LENGTH: 2095  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-249

Query Match 12.7%; Score 185.6; DB 16; Length 2095;  
Best Local Similarity 53.7%; Pred. No. 1.2e-46;  
Matches 464; Conservative 0; Mismatches 379; Indels 21; Gaps 3;  
Qy 190 TACCAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGAGGTTCA 249  
Db 286 TACCTGTTCAGCTGGCTTTTGATTTTCCCGCCACCCAGGAGATGGTGTATCAGGCCACC 345  
Qy 250 GTGCAGCCCATCTTAAGGCACCTTGCTGGAGGCGCAGAAATCCAGTGTGCTTGCCTATGGA 309  
Db 346 ACCAAGAGCCTCATCGAGGGCGTCTATCAGGCTACAAATGCCACTGTCTTTGCTATGCG 405

QY 310 CCACAGAGCTGGGAAGACGACACATGCTGGGAGCCCGCAGAGCAACCTGGGGTGATC 369  
 Db 406 CCACAGAGCTGGGAAGACCTACACATGCTGGGACACAGACAGGAGCTGGGCTAT 465  
 QY 370 CGCGGGCTCTCATGACCTCTGAGCTCACAGGAGGAGGGTGGCGAGGGCGGCCA 429  
 Db 466 GTTCAGACCTCTCAACGACCTCTTCGTGCTCATCGAGGAGACAGCAATGA-----C 516  
 QY 430 TGGGCGCTTCTGTCTACCATGCTTTACCTAGAGATCTACAGGAGAAGTATAGACCTC 489  
 Db 517 ATGGAGTATGAGTCTCTCCATGCTTACCTGGAGATCTACATGAGATGATCGGGACCTG 576  
 QY 490 CTGGACCTGCTTGGGAGACCTGTPATCCGAGAAGACTCGGGGGGATATCTGATT 549  
 Db 577 CTGAACCCCTCTCCCTGGCTTACCTGGAGCTCGGGAGGACTTAAGGGGGTGTATCAGGTG 636  
 QY 550 CGGGTCTCTCCAGAGCCCATCATGCTAGCTTTGCTGATTCTTGGCGGCATCTCTCGCCA 609  
 Db 637 GCCGGCATCACGAGTCTCCACCATCATGCGAGGAGATCATGAGTGTGATGAAG 696  
 QY 610 GCCAGTCGAATCGGACTGTAGAGCCACCGGCTCAACAGCGCTCTCTCCGCGATCAT 669  
 Db 697 GGCAACCGGAGAGGACCGAGGACCCAGCGGCCCAACAGAGCTCTCTCCGCTCCAC 756  
 QY 670 GCTGTGCTCTGCTCAAGTGGACCGGAGGAGCTTTGGCCCCATTTCCGCA-----G 723  
 Db 757 GCGGTACTGAGGTGACCGGTGCGGCCAGCGGAGCGGCTCAAGAACATCTTGAGGAGGTG 816  
 QY 724 CGAGAGGAAATCTTACCTGATTGCTGGGTGAGAGCAACACCGGGCGCACAGCG 783  
 Db 817 CGGAGGGCGCTGTTTCATGCTGAGCTGGTGGCTCAGAGCGGCGCTCGCAGACAG 876  
 QY 784 AACAGGGCTTCCGCTTAAAGAGAGTGGAGCATCAACACCTCTCTGTTGCTTGGGC 843  
 Db 877 AATCGTGGGAGCGTATGAAGAGGGGGGCCCAATCAACCGCTCACTGCTGGCACTGGC 936  
 QY 844 AAGTGTAGATGGCTG-----AATCAGGGCTCTCTGCTGCTTATCGGACAGC 897  
 Db 937 AACTGCATCAACGCGCTTGGCGACAGGGTAGCAACAGTATCACTATCGGACAGC 996  
 QY 898 AAGCTCACTCGCTTATGAGGAGTCTCTGGGTGGCTCAGCCACAGTATCTTATGTC 957  
 Db 997 AAGCTACCGGCTCTGAGGAGTCTCTGGAGGAAACAGCGCGCAGTATGATCGCT 1056  
 QY 958 AACATTGCGCTGAGAGAGCTTCTACTAGACAGCTCTCCGCACTCACTTGTGCTGC 1017  
 Db 1057 CACATCAGTCTCTGAGCAGTGTCTTGGAGAGTCCCGAAGAGTCCCGAAGAGTCCGCGGC 1116  
 QY 1018 AGGTCCAGGAGGTGATCAATCGG 1041  
 Db 1117 CGGCGCAAGAACATTAAGACTAGG 1140

RESULT 15

US-10-287-226-301  
 ; Sequence 301, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agee, Michele L.,  
 ; APPLICANT: Alsbrook, John P.,  
 ; APPLICANT: Berghs, Constance,  
 ; APPLICANT: Boldog, Ferenc,  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.,  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Eisinger, Shlomit R.,  
 ; APPLICANT: Eisen, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khrantsov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malyankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Millet, Charles E.,  
 ; APPLICANT: Millet, Isabelle,  
 ; APPLICANT: Oot, Chean Eng,  
 ; APPLICANT: Ort, Tatiana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastelli, Luca,  
 ; APPLICANT: Rieser, Daniel K.,  
 ; APPLICANT: Rothenberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,  
 ; APPLICANT: Spyttek, Kimberley A.,  
 ; APPLICANT: Taupier, Jr., Raymond J.,  
 ; APPLICANT: Vernet, Corine A.M.,  
 ; APPLICANT: Zerhusen, Bryan D.,  
 ; APPLICANT: Zhong, Mei  
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: 21402-480C  
 ; CURRENT APPLICATION NUMBER: US/10/287,226  
 ; CURRENT FILING DATE: 2002-11-04  
 ; PRIOR APPLICATION NUMBER: 60/334,421  
 ; PRIOR FILING DATE: 2001-11-30  
 ; PRIOR APPLICATION NUMBER: 60/354,392  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: 60/360,148  
 ; PRIOR FILING DATE: 2002-02-27  
 ; PRIOR APPLICATION NUMBER: 60/364,000  
 ; PRIOR FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: 60/404,821  
 ; PRIOR FILING DATE: 2002-08-20  
 ; PRIOR APPLICATION NUMBER: 60/334,526  
 ; PRIOR FILING DATE: 2001-11-30  
 ; PRIOR APPLICATION NUMBER: 60/354,409  
 ; PRIOR FILING DATE: 2002-02-04  
 ; PRIOR APPLICATION NUMBER: 60/364,227  
 ; PRIOR FILING DATE: 2002-03-13  
 ; PRIOR APPLICATION NUMBER: 60/334,027  
 ; PRIOR FILING DATE: 2001-11-28  
 ; PRIOR APPLICATION NUMBER: 60/331,641  
 ; PRIOR FILING DATE: 2001-11-20  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 673  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 301  
 ; LENGTH: 3374  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (7)..(3351)  
 ; US-10-287-226-301  
 Query Match 10.7%; Score 156.8; DB 16; Length 3374;  
 Best Local Similarity 54.1%; Pred. No. 1.3e-37;  
 Matches 395; Conservative 0; Mismatches 317; Indels 18; Gaps 3;  
 QY 321 TGGGAAGACGACACAAATGCTGGGAGCCCGCAGAGCAACCTGGGGTGATCCCGGGGCTCT 380  
 Db 612 TGGGAAGACCTACACCATGCTGGGACAGACAGGAGCTGATCTATGTTTCAGACCT 671  
 QY 381 CATGGACCTCTTCAGCTCACAAGGGAGGAGGTGCCGAGGCGCCGACCATGGCCCTTC 440  
 Db 672 CACAGACCTCTTCGTCGTCATCGAGGAGACCCAGCAATGA-----CATGGATATGA 722  
 QY 441 TGTCCACATGCTTACCTAGAGATCTTACGAGGAGAGGTATTAGACCTCTCTGGACCTTC 500  
 Db 723 GGTTCTCATGTCTTACTCTGAGATCTTACAATGAGATGATCCCGGAGCTGCTGAACCCCTC 782



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:57:04 ; Search time 132.833 Seconds  
(without alignments)  
7833.865 Million cell updates/sec

Title: US-10-797-893-7

Perfect score: 1464

Sequence: 1 atgggtcgctgtcggtgaag.....accattgtccacaatgtga 1464

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1464	100.0	1464	3	US-09-724-224-7
2	1464	100.0	1464	4	US-10-093-317-7
3	1461	99.8	1538	3	US-09-724-224-3
4	1461	99.8	1538	4	US-10-093-317-3
5	1458	99.6	2097	4	US-09-595-684B-34
6	1038.4	70.9	1041	3	US-09-724-224-5
7	1038.4	70.9	1041	4	US-10-093-317-5
8	1035.4	70.7	1115	3	US-09-724-224-1
9	1035.4	70.7	1115	4	US-10-093-317-1
10	138.6	9.5	4108	4	US-09-883-096-1
11	136.6	9.3	1152	4	US-09-883-096-6
12	128.8	8.8	897	3	US-09-621-233-1
13	128.8	8.8	897	3	US-09-724-224-1
14	128.8	8.8	897	3	US-09-724-516-1
15	128.8	8.8	897	4	US-10-090-695-1
16	126.6	8.6	1026	3	US-09-641-806-1
17	126.6	8.6	1026	4	US-09-723-129-1
18	126.6	8.6	1026	4	US-09-722-862-1
19	124.2	8.5	1011	3	US-09-641-806-3
20	124.2	8.5	1011	4	US-09-723-129-3
21	124.2	8.5	1011	4	US-09-722-862-3
22	121.4	8.3	1014	4	US-09-883-096-3
23	101	6.9	3258	4	US-09-799-451-586
24	100	6.8	3572	2	US-08-713-815A-2
25	87.6	6.0	4014	3	US-09-541-782-1
26	87.6	6.0	4014	4	US-09-723-820-1
27	87.6	6.0	4014	4	US-10-270-085-1

28	79.6	5.4	1032	4	US-09-967-908A-7	Sequence 7, Appli
29	79.6	5.4	1032	4	US-10-159-151-7	Sequence 7, Appli
30	79.6	5.4	1122	4	US-09-967-908A-3	Sequence 3, Appli
31	79.6	5.4	1122	4	US-10-159-151-3	Sequence 3, Appli
32	79.4	5.4	3930	3	US-09-162-373-2	Sequence 2, Appli
33	79.4	5.4	3930	3	US-09-467-946-2	Sequence 2, Appli
34	79.2	5.4	1092	4	US-09-967-908A-9	Sequence 9, Appli
35	79.2	5.4	1092	4	US-10-159-151-9	Sequence 9, Appli
36	79.2	5.4	1146	4	US-09-967-908A-5	Sequence 5, Appli
37	79.2	5.4	1146	4	US-10-159-151-5	Sequence 5, Appli
38	79.2	5.4	6409	4	US-09-967-908A-1	Sequence 1, Appli
39	79.2	5.4	6409	4	US-10-159-151-1	Sequence 1, Appli
40	78.4	5.4	1065	3	US-09-724-511-3	Sequence 3, Appli
41	78.4	5.4	1065	4	US-09-723-097-3	Sequence 3, Appli
42	78.4	5.4	1065	4	US-09-632-344-3	Sequence 3, Appli
43	78.4	5.4	1217	3	US-09-724-511-1	Sequence 1, Appli
44	78.4	5.4	1217	4	US-09-723-097-1	Sequence 1, Appli
45	78.4	5.4	1217	4	US-09-632-344-1	Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-09-724-224-7  
; Sequence 7, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724, 224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597, 292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1464  
; TYPE: DNA  
; ORGANISM: Human  
US-09-724-224-7

Query Match		100.0%	Score 1464;	DB 3;	Length 1464;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1464;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTGCGCTGTCGGCTAACCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGGTA	60		
DB	1	ATGGGTGCGCTGTCGGCTAACCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGGTA	60		
QY	61	AGGGTGGCTGTCGGACTCGGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCC	120		
DB	61	AGGGTGGCTGTCGGACTCGGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCC	120		
QY	121	TGTGTGGGGGCAATGGACAGCTGCTCTAGAGATTCTTAACAGGAAACCAACAGAG	180		
DB	121	TGTGTGGGGGCAATGGACAGCTGCTCTAGAGATTCTTAACAGGAAACCAACAGAG	180		
QY	181	ACTCTCAATACCAAGTTTGAATGCTTATGGGAGAGGAGTACTCAGCAGGACATCTAT	240		
DB	181	ACTCTCAATACCAAGTTTGAATGCTTATGGGAGAGGAGTACTCAGCAGGACATCTAT	240		
QY	241	GCAGGTTTCAGTCAGGCCCATCTTAAGCACTTGTGGAAGGCGAGATGCCAGTGTGTT	300		
DB	241	GCAGGTTTCAGTCAGGCCCATCTTAAGCACTTGTGGAAGGCGAGATGCCAGTGTGTT	300		
QY	301	GCCTATGGACCAACAGAGCTGGGAAGACGCAACAAATGTGGGCAAGCCAGAGCAACT	360		
DB	301	GCCTATGGACCAACAGAGCTGGGAAGACGCAACAAATGTGGGCAAGCCAGAGCAACT	360		
QY	361	GGGGTGATCCCGGGGCTCTCATGGACCTCTCATGAGTCTACAGGGAGGAGGGTCCGAG	420		

361 GGGGTGATCCGGGGCTCTCATGACCTCTCGAGCTCACAGGGAGGGTCCGAG 420  
 421 GCGCGGCCATGGGCCCTTTCTGTACATGTCTTACCTAGAGATCTACAGGAGAGTA 480  
 421 GCGCGGCCATGGGCCCTTTCTGTACATGTCTTACCTAGAGATCTACAGGAGAGTA 480  
 481 TTAGACCTCTGGACCTCTTCCGAGACCTGGTAATCCGAGAGACTGCCGGGGAT 540  
 481 TTAGACCTCTGGACCTCTTCCGAGACCTGGTAATCCGAGAGACTGCCGGGGAT 540  
 541 ATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600  
 541 ATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600  
 601 TTCTGCCAGCCAGTCGAATCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660  
 601 TTCTGCCAGCCAGTCGAATCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660  
 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGACCGAGCGGAGCGTTTGGCCCATTTCCG 720  
 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGACCGAGCGGAGCGTTTGGCCCATTTCCG 720  
 721 CAGCGAGAGGAACTCTACCTGATTTGACTTGGCTGGGTCAAGAGCAACCGGCGACA 780  
 721 CAGCGAGAGGAACTCTACCTGATTTGACTTGGCTGGGTCAAGAGCAACCGGCGACA 780  
 781 GGCACCAAGGGCTTCCGGTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTG 840  
 781 GGCACCAAGGGCTTCCGGTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTG 840  
 841 GGCACCAAGGGCTTCCGGTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTG 900  
 841 GGCACCAAGGGCTTCCGGTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTG 900  
 901 CTCACTCGCCTATGTCAGGACTCTCTGGGTGGCTCAGCCCAACAGTATCTTATGCGAAC 960  
 901 CTCACTCGCCTATGTCAGGACTCTCTGGGTGGCTCAGCCCAACAGTATCTTATGCGAAC 960  
 961 ATTGCCCTTGAGAGAGCTTTACCTAGACACAGTCTCCGACTCAACTTTGTCGAGG 1020  
 961 ATTGCCCTTGAGAGAGCTTTACCTAGACACAGTCTCCGACTCAACTTTGTCGAGG 1020  
 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCTCTATGCTTTGGGA 1080  
 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCTCTATGCTTTGGGA 1080  
 1081 CTTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGCCGAGGC 1140  
 1081 CTTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGCCGAGGC 1140  
 1141 CTTGAGAGAGAGAGATTTGGAGCCCTGAGCCCATGGAGCTTCAGCCTCTGCTCCGAG 1200  
 1141 CTTGAGAGAGAGAGATTTGGAGCCCTGAGCCCATGGAGCTTCAGCCTCTGCTCCGAG 1200  
 1201 AACTCAGCCCTCAGAGAGCTTAAAGAGATGAGCCATGAGCTTGGAGCGCTCTCTC 1260  
 1201 AACTCAGCCCTCAGAGAGCTTAAAGAGATGAGCCATGAGCTTGGAGCGCTCTCTC 1260  
 1261 AGCTTGGACCTGTGCTCTCCAGGGGAGGAGCGAGCGCCCTCTGTTGAGTACCCCA 1320  
 1261 AGCTTGGACCTGTGCTCTCCAGGGGAGGAGCGAGCGCCCTCTGTTGAGTACCCCA 1320  
 1321 AAGCGAGAGCGGATGCTGCTTAATGAGACAGTATGAGAGAGGACCTAGAGATTGAGG 1380  
 1321 AAGCGAGAGCGGATGCTGCTTAATGAGACAGTATGAGAGAGGACCTAGAGATTGAGG 1380  
 1381 CTTAAGACGAGCAAAAGAACTTGGAGCCCAAGATGTTGGCCAGAGGCTGAGGAAAG 1440  
 1381 CTTAAGACGAGCAAAAGAACTTGGAGCCCAAGATGTTGGCCAGAGGCTGAGGAAAG 1440  
 1441 GAGAACCATTTGCCCAATGTGA 1464

Db 1441 GAGAACCATTTGTCCCAATGTGA 1464  
 RESULT 2  
 US-10-093-317-7  
 ; Sequence 7, Application US/10093317  
 ; Patent No. 6762043  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/093,317  
 ; CURRENT FILING DATE: 2002-03-06  
 ; PRIOR APPLICATION NUMBER: 09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 1464  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-093-317-7  
 Query Match 100.0%; Score 1464; DB 4; Length 1464;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGGTGCTGTGCGGCTAAGCAAGATTTGGAGCTACTCGTCTGCCACTCCAGCTCGCGTA 60  
 Db 1 ATGGGTGCTGTGCGGCTAAGCAAGATTTGGAGCTACTCGTCTGCCACTCCAGCTCGCGTA 60  
 QY 61 AGGTGGCTGTGCGGCTAAGCAAGATTTGGAGCTACTCGTCTGCCACTCCAGCTCGCGTA 120  
 Db 61 AGGTGGCTGTGCGGCTAAGCAAGATTTGGAGCTACTCGTCTGCCACTCCAGCTCGCGTA 120  
 QY 121 TGTGTGCGGGGATGAGACAGCTGCTCTTAGAGATTTGCTAACTGGAGGAACACCAAGAG 180  
 Db 121 TGTGTGCGGGGATGAGACAGCTGCTCTTAGAGATTTGCTAACTGGAGGAACACCAAGAG 180  
 QY 181 ACTCTCAATATACAGTTTGTATGCTCTTATGGGGAGAGGACTCTCAGCAGACATCTAT 240  
 Db 181 ACTCTCAATATACAGTTTGTATGCTCTTATGGGGAGAGGACTCTCAGCAGACATCTAT 240  
 QY 241 CGAGGTTTCACTGAGCCCATCTTAAAGGCACTTGTCTGGAAGGCGAGAAATGCCAGTGTCT 300  
 Db 241 CGAGGTTTCACTGAGCCCATCTTAAAGGCACTTGTCTGGAAGGCGAGAAATGCCAGTGTCT 300  
 QY 301 GCCTATGGACCCACAGGAGCTGGGAGACGCAACAATGCTGGGAGCCAGAGCAACCT 360  
 Db 301 GCCTATGGACCCACAGGAGCTGGGAGACGCAACAATGCTGGGAGCCAGAGCAACCT 360  
 QY 361 GGGGTGATCCCGGGGCTCTCATGACCTCTTGCAGCTCACAAGGGAGAGGGTCCGAG 420  
 Db 361 GGGGTGATCCCGGGGCTCTCATGACCTCTTGCAGCTCACAAGGGAGAGGGTCCGAG 420  
 QY 421 GSCCGCCATGSGCCCTTTCTGTACCATGCTTACCTAGAGATCTTACAGGAGAAAGTA 480  
 Db 421 GSCCGCCATGSGCCCTTTCTGTACCATGCTTACCTAGAGATCTTACAGGAGAAAGTA 480  
 QY 481 TTAGACCTCTCGGACCTCTTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGGGAT 540  
 Db 481 TTAGACCTCTCGGACCTCTTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGGGAT 540  
 QY 541 ATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600  
 Db 541 ATCTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600  
 QY 601 TTCTGCCAGCCAGTCGAATCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660  
 Db 601 TTCTGCCAGCCAGTCGAATCGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660  
 QY 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGACCGAGCGGAGCGTTTGGCCCATTTCCG 720

Db 661 CGCAGTCATGCTGCTCTCTGCTCAAGGTGGACCAAGCGGGAACGTTTGGCCCCCATTTGCG 720  
Qy 721 CAGCGAGAGGAAACTCTACTGATTCAGTTTGGCTGGCTCAGAGGACAAACCGGCGACA 780  
Db 721 CAGCGAGAGGAAACTCTACTGATTCAGTTTGGCTGGCTCAGAGGACAAACCGGCGACA 780  
Qy 781 GGCACAAAGGCGCTTGGCTTAAAGAGAGAGTGAGGCATCAACACCTCCCTGTTTGTCTG 840  
Db 781 GGCACAAAGGCGCTTGGCTTAAAGAGAGAGTGAGGCATCAACACCTCCCTGTTTGTCTG 840  
Qy 841 GGCACAAAGGCGCTTGGCTTAAAGAGAGAGTGAGGCATCAACACCTCCCTGTTTGTCTG 900  
Db 841 GGCACAAAGGCGCTTGGCTTAAAGAGAGAGTGAGGCATCAACACCTCCCTGTTTGTCTG 900  
Qy 901 CTCACTCGCTTATGTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTTGCCAAC 960  
Db 901 CTCACTCGCTTATGTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTATTTGCCAAC 960  
Qy 961 ATTGCCCTGAGAGACGCTTCTAAGACAGCTCTCCGACCTCAACTTTTGTGCCAGG 1020  
Db 961 ATTGCCCTGAGAGACGCTTCTAAGACAGCTCTCCGACCTCAACTTTTGTGCCAGG 1020  
Qy 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGAGCCTCATGCTTGGGA 1080  
Db 1021 TCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGAGCCTCATGCTTGGGA 1080  
Qy 1081 CCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140  
Db 1081 CCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140  
Qy 1141 CCTGAGGAAGAGAGATTTGGAGGCTTGAGCCCATGGCAGCTCCAGCCTCTGCTCCAG 1200  
Db 1141 CCTGAGGAAGAGAGATTTGGAGGCTTGAGCCCATGGCAGCTCCAGCCTCTGCTCCAG 1200  
Qy 1201 AAACCTAGCCCCCTACAGAGCTAAGCAGATGAGACCGGCGCATGCTGGAGCGCTCTC 1260  
Db 1201 AAACCTAGCCCCCTACAGAGCTAAGCAGATGAGACCGGCGCATGCTGGAGCGCTCTC 1260  
Qy 1261 AGCTTGACCGCTGCTGCTGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
Db 1261 AGCTTGACCGCTGCTGCTGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
Qy 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1380  
Db 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1380  
Qy 1381 CTTAAGACGAAAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCAGAGAGGCTGAGGAAAG 1440  
Db 1381 CTTAAGACGAAAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCAGAGAGGCTGAGGAAAG 1440  
Qy 1441 GAGAACCAATGCTCCCAATGTGA 1464  
Db 1441 GAGAACCAATGCTCCCAATGTGA 1464

## RESULT 3

US-09-724-224-3  
; Sequence 3, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724,224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1538

; TYPE: DNA  
; ORGANISM: Human  
US-09-724-224-3

Query Match 99.8%; Score 1461; DB 3; Length 1538;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGGTAAGG 63  
Db 78 GGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGGTAAGG 137  
Qy 64 GTGCTGTGGACTGCGGCCATTTGTGGATGGAACAGCGGAGCAAGTGAATCCCTCTGT 123  
Db 138 GTGCTGTGGACTGCGGCCATTTGTGGATGGAACAGCGGAGCAAGTGAATCCCTCTGT 197  
Qy 124 GTGCGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGAGACT 183  
Db 198 GTGCGGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGAGACT 257  
Qy 184 CTCAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243  
Db 258 CTCAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317  
Qy 244 GGTTCAGTCAGAGCCCATCTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTTGC 303  
Db 318 GGTTCAGTCAGAGCCCATCTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTTGC 377  
Qy 304 TATGGAACCAAGAGCTGGGAAGACGCAACAATGTGGGACGCCAGCAGCAACCTGGG 363  
Db 378 TATGGAACCAAGAGCTGGGAAGACGCAACAATGTGGGACGCCAGCAGCAACCTGGG 437  
Qy 364 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAACAGGAGGAGGAGGAGGAGGAGG 423  
Db 438 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAACAGGAGGAGGAGGAGGAGGAGG 497  
Qy 424 CGGCAATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACACAGGAGAAATATTA 483  
Db 498 CGGCAATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACACAGGAGAAATATTA 557  
Qy 484 GACCTCTGAGACCTTGTCTGGGAGACCTGGTAAATCCGAGAAGACTGCGGGGGAATATC 543  
Db 558 GACCTCTGAGACCTTGTCTGGGAGACCTGGTAAATCCGAGAAGACTGCGGGGGAATATC 617  
Qy 544 CTGATTCGGGCTCTTCCCAAGAGCCCATCAGTAGCTTCTGATGATTTGAGCGGACCTTC 603  
Db 618 CTGATTCGGGCTCTTCCCAAGAGCCCATCAGTAGCTTCTGATGATTTGAGCGGACCTTC 677  
Qy 604 CTGCGACCAAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCTCTCTCCGC 663  
Db 678 CTGCGACCAAGTGGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCTCTCTCCGC 737  
Qy 664 AGTCATGCTGTGCTCTGCTGTCAGAGGAGCAAGGAGGAACTGTTGGGCCCCCATTTGCGCAG 723  
Db 738 AGTCATGCTGTGCTCTGCTGTCAGAGGAGCAAGGAGGAACTGTTGGGCCCCCATTTGCGCAG 797  
Qy 724 CGAGAGGGAATCTCTACCTGATTTGATCTGGCTGGGTTCAGGAGCAACCGGCGCACAGGC 783  
Db 798 CGAGAGGGAATCTCTACCTGATTTGATCTGGCTGGGTTCAGGAGCAACCGGCGCACAGGC 857  
Qy 784 AACAGGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 843  
Db 858 AACAGGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC 917  
Qy 844 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGGAAGCAAGCTC 903  
Db 918 AAAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGGAAGCAAGCTC 977  
Qy 904 ACTGCGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAACATT 963  
Db 978 ACTGCGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTATCTTATTCGCAACATT 1037  
Qy 964 GCCCTTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCTCAACTTTTGTGCGCAGGTCC 1023





```
Db 1338 TTGGACCGCTGCTTGGCTTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397
Qy 1324 CGAGAGCGGATGGTCTAATCAAGACAGTGAAGAAGAGGACCTAGAGATTGAGAGGCTT 1383
Db 1398 CGAGAGCGGATGGTCTAATCGAGACAGTGAAGAAGAGGACCTAGAGATTGAGAGGCTT 1457
Qy 1384 AAGACGAAGCAAAAAGAACTCGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1443
Db 1458 AAGACGAAGCAAAAAGAACTCGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1517
Qy 1444 AACCAATTGTCCCAATGTGA 1464
Db 1518 AACCAATTGTCCCAATGTGA 1538

RESULT 5
US-09-595-684B-34
; Sequence 34, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cyto36
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-34

Query Match 99.6%; Score 1458; DB 4; Length 2097;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAGG 63
Db 100 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAGG 159
Qy 64 GTGGCTGTGGACTGCGGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 123
Db 160 GTGGCTGTGGACTGCGGCCATTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 219
Qy 124 GTGGGGGCAATGACAGCTGCTCTAGAGATTGCTAACTGGAGAAACCAACAGGAGACT 183
Db 220 GTGGGGGCAATGACAGCTGCTCTAGAGATTGCTAACTGGAGAAACCAACAGGAGACT 279
Qy 184 CTCAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
Db 280 CTCAAATACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 339
Qy 244 GGTTCAGTGCAGCCCACTTCAAGGCACTTGTGGAAGGCGCAAGATGCCAGTGTGCTGCC 303
Db 340 GGTTCAGTGCAGCCCACTTCAAGGCACTTGTGGAAGGCGCAAGATGCCAGTGTGCTGCC 399
Qy 304 TATGACCCCAAGAGCTGGGAAGACGCAACAATGCTGGGCAAGCCAGAGCAACCTGGG 363
Db 400 TATGACCCCAAGAGCTGGGAAGACGCAACAATGCTGGGCAAGCCAGAGCAACCTGGG 459
Qy 364 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAAGAGGAGGAGGCTGCGGAGGC 423
Db 460 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAAGAGGAGGAGGCTGCGGAGGC 519
```

```
Qy 424 CGGCATGGGGCCCTTTCTGTCAACCATGCTTACCTAGAGATCTTACCAGGAGAGCTATTATTA 483
Db 520 CGGCATGGGGCCCTTTCTGTCAACCATGCTTACCTAGAGATCTTACCAGGAGAGCTATTATTA 579
Qy 484 GACCTCTCGGACCTCTGCTTCCGAGAGACTGCTGTAATCCGAGAGAGCTGCGGGGGGAATATC 543
Db 580 GACCTCTCGGACCTCTGCTTCCGAGAGACTGCTGTAATCCGAGAGAGCTGCGGGGGGAATATC 639
Qy 544 CTGATTCGGGGTCTCTCCCAAGCCCATCAGTAGCTTGGCTGATTTTGGGGGACATTC 603
Db 640 CTGATTCGGGGTCTCTCCCAAGCCCATCAGTAGCTTGGCTGATTTTGGGGGACATTC 699
Qy 604 CTGCCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCG 663
Db 700 CTGCCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCG 759
Qy 664 AGTCATGCTGTCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGGCCCAATTTGCCGAG 723
Db 760 AGTCATGCTGTCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGGCCCAATTTGCCGAG 819
Qy 724 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTCAAGAGGACAAACCGGGGACAGGC 783
Db 820 CGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTCAAGAGGACAAACCGGGGACAGGC 879
Qy 784 AACAAAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 843
Db 880 AACAAAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 939
Qy 844 AAAGTGTGATGCTGCTGAATCAGGGCTCTCTGCTGATCTTATCGGACAGCAAGCTC 903
Db 940 AAAGTGTGATGCTGCTGAATCAGGGCTCTCTGCTGATCTTATCGGACAGCAAGCTC 999
Qy 904 ACTCGCTATTCAGGAGCTCTCTGGGTGGCTCAGCCCCACAGTATCTTATTCGAACATT 963
Db 1000 ACTCGCTATTCAGGAGCTCTCTGGGTGGCTCAGCCCCACAGTATCTTATTCGAACATT 1059
Qy 964 GCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTTCAACTTTGTCTGCCAGTCC 1023
Db 1060 GCCCTGAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTTCAACTTTGTCTGCCAGTCC 1119
Qy 1024 AAGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1083
Db 1120 AAGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1179
Qy 1084 GTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT 1143
Db 1180 GTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT 1239
Qy 1144 GAGGAAGAGAGATTTGGGAGCCCTGAGCCCATGSCAGCTCCAGGCTCTGCTCTCCAGAAA 1203
Db 1240 GAGGAAGAGAGATTTGGGAGCCCTGAGCCCATGSCAGCTCCAGGCTCTGCTCTCCAGAAA 1299
Qy 1204 CTGAGCCCCCTCAGAAAGCTTAAGCAGCATCGAACCCGGCCATGTGGAGCGCTCTCTCAGC 1263
Db 1300 CTGAGCCCCCTCAGAAAGCTTAAGCAGCATGGAACCCGGCCATGTGGAGCGCTCTCTCAGC 1359
Qy 1264 TTGGAACGCTGTCTGCTCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323
Db 1360 TTGGAACGCTGTCTGCTCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1419
Qy 1324 CGAGAGCGGATGGTGTCTAATGAAGACAGTGAAGAAGAGGACCTTAGAGATTGAGAGGCTT 1383
Db 1420 CGAGAGCGGATGGTGTCTAATGAAGACAGTGAAGAAGAGGACCTTAGAGATTGAGAGGCTT 1479
Qy 1384 AAGACGAAGCAAAAAGAACTGGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1443
Db 1480 AAGACGAAGCAAAAAGAACTGGAGGCCCAAGATGTTGGCCCCAGAAAGGCTGAGGAAAAGGAG 1539
Qy 1444 AACCAATTGTCCCAATG 1461
Db 1540 AACCAATTGTCCCAATG 1557
```

```

RESULT 6
US-09-724-224-5
; Sequence 5, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-5

Query Match          70.9%; Score 1038.4; DB 3; Length 1041;
Best Local Similarity 99.9%; Pred. No. 8.8e-293;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTCTGCTCGTCCACCTCCAGTCCGCTA 60
DB 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTCTGCTCGTCCACCTCCAGTCCGCTA 60

QY 61 AGGGTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
DB 61 AGGGTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120

QY 121 TGTGTGGGGGATGACAGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180
DB 121 TGTGTGGGGGATGACAGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180

QY 181 ACTCTCAAAATACCAATTTGATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTAT 240
DB 181 ACTCTCAAAATACCAATTTGATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTAT 240

QY 241 GCAGGTTCAAGTCCAGCCATCCAGGAGTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 300
DB 241 GCAGGTTCAAGTCCAGCCATCCAGGAGTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 300

QY 301 GCCTATGACCCACAGGAGTGGGAAGACGACACAAATGCTGGGGAGCCCGAGCAACCT 360
DB 301 GCCTATGACCCACAGGAGTGGGAAGACGACACAAATGCTGGGGAGCCCGAGCAACCT 360

QY 361 GGGGTGATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420
DB 361 GGGGTGATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420

QY 421 GCGCGGCATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGT 480
DB 421 GCGCGGCATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGT 480

QY 481 TTAGACCTCTCTGAGCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAAGTCCCGGGGAAT 540
DB 481 TTAGACCTCTCTGAGCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAAGTCCCGGGGAAT 540

QY 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGCTGATTTTGGAGCGCAC 600
DB 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGCTGATTTTGGAGCGCAC 600

QY 601 TTCCTGCCAGCAGTCCGAATTCGAGTCTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
DB 601 TTCCTGCCAGCAGTCCGAATTCGAGTCTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660

QY 661 CGCAGTCAATGCTGCTCTCTGCTCAAGTGAACAGCGGGAAACGTTTGGGCCCAATTCGC 720
DB 661 CGCAGTCAATGCTGCTCTCTGCTCAAGTGAACAGCGGGAAACGTTTGGGCCCAATTCGC 720

```

```

QY 721 CAGCGAGAGGAAACTCTACCTGATTGACTTGGCTGGGTTCAGAGGACACCGCGCAC 780
DB 721 CAGCGAGAGGAAACTCTACCTGATTGACTTGGCTGGGTTCAGAGGACACCGCGCAC 780

QY 781 GGCACCAAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTG 840
DB 781 GGCACCAAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTG 840

QY 841 GGCACCAAGTGTAGATGCGCTGAATTCAGGGCTCCCTGCTGTACTCTTATTCGGGACAGCAAG 900
DB 841 GGCACCAAGTGTAGATGCGCTGAATTCAGGGCTCCCTGCTGTACTCTTATTCGGGACAGCAAG 900

QY 901 CTCACTGCGCTATTCAGAGGACTCTCTGGGTGGGTTCAGCCACACAGTATCTTATTCGCAAC 960
DB 901 CTCACTGCGCTATTCAGAGGACTCTCTGGGTGGGTTCAGCCACACAGTATCTTATTCGCAAC 960

QY 961 ATTGCGCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020
DB 961 ATTGCGCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020

QY 1021 TCCAAAGGAGTGAATCAATCG 1040
DB 1021 TCCAAAGGAGTGAATCAATTTG 1040

RESULT 7
US-10-093-317-5
; Sequence 5, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-5

Query Match          70.9%; Score 1038.4; DB 4; Length 1041;
Best Local Similarity 99.9%; Pred. No. 8.8e-293;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTCTGCTCGTCCACCTCCAGTCCGCTA 60
DB 1 ATGGGTGCGTGTGCGGCTAAGCAAGATTGGAGTCTCTGCTCGTCCACCTCCAGTCCGCTA 60

QY 61 AGGGTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
DB 61 AGGGTGGCTGTGCGACTGCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120

QY 121 TGTGTGGGGGATGACAGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180
DB 121 TGTGTGGGGGATGACAGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180

QY 181 ACTCTCAAAATACCAATTTGATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTAT 240
DB 181 ACTCTCAAAATACCAATTTGATGCTTCTATGGGGAGAGAGTACTCAGCAGACATCTAT 240

QY 241 GCAGGTTCAAGTCCAGCCATCCAGGAGTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 300
DB 241 GCAGGTTCAAGTCCAGCCATCCAGGAGTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 300

QY 301 GCCTATGACCCACAGGAGTGGGAAGACGACACAAATGCTGGGGAGCCCGAGCAACCT 360
DB 301 GCCTATGACCCACAGGAGTGGGAAGACGACACAAATGCTGGGGAGCCCGAGCAACCT 360

QY 361 GGGGTGATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420
DB 361 GGGGTGATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAAGAGGAGAGGGTGGCGAG 420

QY 421 GCGCGGCATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGT 480
DB 421 GCGCGGCATGGGCGCTTCTGTCACCAATGCTTACCTAGAGATCTACAGGAGAGAGT 480

QY 481 TTAGACCTCTCTGAGCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAAGTCCCGGGGAAT 540
DB 481 TTAGACCTCTCTGAGCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAAGTCCCGGGGAAT 540

QY 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGCTGATTTTGGAGCGCAC 600
DB 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGCTGATTTTGGAGCGCAC 600

QY 601 TTCCTGCCAGCAGTCCGAATTCGAGTCTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660
DB 601 TTCCTGCCAGCAGTCCGAATTCGAGTCTAGAGCCACCCGGCTCAACAGCGCTCTCTCC 660

QY 661 CGCAGTCAATGCTGCTCTCTGCTCAAGTGAACAGCGGGAAACGTTTGGGCCCAATTCGC 720
DB 661 CGCAGTCAATGCTGCTCTCTGCTCAAGTGAACAGCGGGAAACGTTTGGGCCCAATTCGC 720

```

```

Db 301 GCCTATGACCAACGAGAGCTGGGAAGACGACACAAATGCTGGGAGCCCGCAGAGCAACCT 360
Qy 361 GGGGTGATCCGCGGGCTCTCATGGACCTCTGAGACCTCTGAGCTCAACAAGGAGGAGGGTCCGAG 420
Db 361 GGGGTGATCCGCGGGCTCTCATGGACCTCTGAGACCTCTGAGCTCAACAAGGAGGAGGGTCCGAG 420
Qy 421 GCGCGGCATGGGCGCTTCTGTCACATGTCCTTACCTAGAGATCTACAGAGAGGTA 480
Db 421 GCGCGGCATGGGCGCTTCTGTCACATGTCCTTACCTAGAGATCTACAGAGAGGTA 480
Qy 481 TTAGACCTCTCGGACCTCTGCTTCGGGAGACCTGTAATCCAGAGAGCTGCGGGGGAAT 540
Db 481 TTAGACCTCTCGGACCTCTGCTTCGGGAGACCTGTAATCCAGAGAGCTGCGGGGGAAT 540
Qy 541 ATCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCAC 600
Db 541 ATCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCAC 600
Qy 601 TTCTGCGCAGCAGTCGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGCGTCTCTCC 660
Db 601 TTCTGCGCAGCAGTCGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGCGTCTCTCC 660
Qy 661 CGCAGTCATGCTGTCTCTGCTCAAGTGAACGAGGGAAACGTTTGGGCGGCAATTCGC 720
Db 661 CGCAGTCATGCTGTCTCTGCTCAAGTGAACGAGGGAAACGTTTGGGCGGCAATTCGC 720
Qy 721 CAGCGAGAGGAAACTCTACTGATTCAGTTCAGTGGCTGAGGAGCAACCGCGGCACA 780
Db 721 CAGCGAGAGGAAACTCTACTGATTCAGTTCAGTGGCTGAGGAGCAACCGCGGCACA 780
Qy 781 GCGAAAGAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCTGTTTGTCTG 840
Db 781 GCGAAAGAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCTCTGTTTGTCTG 840
Qy 841 GGCAAGTGTAGATGCGCTGAATCAGGCGCTCTGCTGCTGACCTTATCGGAGCAGCAAG 900
Db 841 GGCAAGTGTAGATGCGCTGAATCAGGCGCTCTGCTGCTGACCTTATCGGAGCAGCAAG 900
Qy 901 CTCACTCGCTTATGACGAGTCTCTGGTGGCTCAGCCACAGTATCTTATGCGCAAC 960
Db 901 CTCACTCGCTTATGACGAGTCTCTGGTGGCTCAGCCACAGTATCTTATGCGCAAC 960
Qy 961 ATTGCGCTGAGAGAGCGTCTTACTAGACACAGTCTCGGACCTCAACTTTGTCGCGAG 1020
Db 961 ATTGCGCTGAGAGAGCGTCTTACTAGACACAGTCTCGGACCTCAACTTTGTCGCGAG 1020
Qy 1021 TCCAAGGAGTGATCAATCG 1040
Db 1021 TCCAAGGAGTGATCAATTG 1040

```

RESULT 8

```

; Sequence 1, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-1

```

Query Match

70.7%; Score 1035.4; DB 3; Length 1115;

```

Best Local Similarity 99.9%; Pred. No. 6.9e-292;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 4 GGTGCTGTGGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 63
Db 78 GGTGCTGTGGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 137
Qy 64 GTGGCTGTGGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 123
Db 138 GTGGCTGTGGGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGCGTAAGG 197
Qy 124 GTGGGCGCATGAGACAGCTGCTCTTAGAGATTCCTAATCTGGAGAAACCAACAGAGACT 183
Db 198 GTGGGCGCATGAGACAGCTGCTCTTAGAGATTCCTAATCTGGAGAAACCAACAGAGACT 257
Qy 184 CTCAAAATACCAATTTGATGCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
Db 258 CTCAAAATACCAATTTGATGCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
Qy 244 GGTTCAGTCAGCCCACTTAAGGCACTTGTCTGGAAGGCGAGAATGCCAGTGTCTTGCC 303
Db 318 GGTTCAGTCAGCCCACTTAAGGCACTTGTCTGGAAGGCGAGAATGCCAGTGTCTTGCC 377
Qy 304 TATGGAACCAAGAGCTGGGAAGACGCAACAATGTGGGAGCCCGCAGAGCAACCTGGG 363
Db 378 TATGGAACCAAGAGCTGGGAAGACGCAACAATGTGGGAGCCCGCAGAGCAACCTGGG 437
Qy 364 GTGATCCCGGGGCTCTCATGGACCTCTGAGGAGTCTCAAGGAGGAGGGTCCGAGGGC 423
Db 438 GTGATCCCGGGGCTCTCATGGACCTCTGAGGAGTCTCAAGGAGGAGGGTCCGAGGGC 497
Qy 424 CGGCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 483
Db 498 CGGCCATGGGCGCTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 557
Qy 484 GACCTCTGAGACCTGCTTCCGAGAGCTGGTAAATCCGAGAACTGCCGGGGAATATC 543
Db 558 GACCTCTGAGACCTGCTTCCGAGAGCTGGTAAATCCGAGAACTGCCGGGGAATATC 617
Qy 544 CTGATTCGGGGTCTCTCCAGAGAGCCATCAGTAGCTTGTGATTTTGGCGGCACTTC 603
Db 618 CTGATTCGGGGTCTCTCCAGAGAGCCATCAGTAGCTTGTGATTTTGGCGGCACTTC 677
Qy 604 CTGCGAGCAGTGGAAATCGGACTGTAGGAGCCACCGGGTCAACAGCGCTCTCCCGC 663
Db 678 CTGCGAGCAGTGGAAATCGGACTGTAGGAGCCACCGGGTCAACAGCGCTCTCCCGC 737
Qy 664 AGTCATGCTGTCTCTGTTCAAGGTGGACCAAGGGGAAACGTTTGGGCGGCAACCTTC 723
Db 738 AGTCATGCTGTCTCTGTTCAAGGTGGACCAAGGGGAAACGTTTGGGCGGCAACCTTC 797
Qy 724 CGAGAGGGAACCTCTACCTGATTTGCTGGGTGGGTGAGGAGCAACCGGCGCAGAGC 783
Db 798 CGAGAGGGAACCTCTACCTGATTTGCTGGGTGGGTGAGGAGCAACCGGCGCAGAGC 857
Qy 784 AACAGGGGCTCTCGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTGGGC 843
Db 858 AACAGGGGCTCTCGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTGGGC 917
Qy 844 AAAGTGTAGATGCGTGAATCAGGGGCTCTCTGTTGATCTTATCGGAGCAGCAAGCTC 903
Db 918 AAAGTGTAGATGCGTGAATCAGGGGCTCTCTGTTGATCTTATCGGAGCAGCAAGCTC 977
Qy 904 ACTCGCTTATGAGGACTCTCTGGGTGGGTGAGCCACAGTATCTTATGCGCAACATT 963
Db 978 ACTCGCTTATGAGGACTCTCTGGGTGGGTGAGCCACAGTATCTTATGCGCAACATT 1037
Qy 964 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTCGCAGTCC 1023
Db 1038 GCCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTCGCAGTCC 1097
Qy 1024 AAGGAGTGATCAATCG 1040

```

```

Db      1098 AAGGAGGTGATCAATTG 1114
Query Match      70.7%; Score 1035.4; DB 4; Length 1115;
Best Local Similarity 99.9%; Pred. No. 6.9e-292;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4  GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTAAGG 63
Db      78  GGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTAAGG 137

QY      64  GTGGCTGTGGACTCGGGCCATTGTGGATGGAACAGCGGAGCAAGTATCCCCCTGT 123
Db      138 GTGGCTGTGGACTCGGGCCATTGTGGATGGAACAGCGGAGCAAGTATCCCCCTGT 197

QY      124 GTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGGAGACT 183
Db      198 GTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGGAGACT 257

QY      184 CTCAATACCAATTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
Db      258 CTCAATACCAATTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317

QY      244 GGTTCAGTCAGGCCATCTTAAGCACTTCTGGAAGGGCAGAAATGCCAGTGTGCTGCC 303
Db      318 GGTTCAGTCAGGCCATCTTAAGCACTTCTGGAAGGGCAGAAATGCCAGTGTGCTGCC 377

QY      304 TATGGACCCACAGGAGCTGGAGACGACACACAATGCTGGGACGCCAGAGCAACCTGGG 363
Db      378 TATGGACCCACAGGAGCTGGAGACGACACACAATGCTGGGACGCCAGAGCAACCTGGG 437

QY      364 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAAGGGAGGAGGTGCCGAGGGC 423
Db      438 GTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAAGGGAGGAGGTGCCGAGGGC 497

QY      424 CGGCCATGGGGCCCTTTCTGTACCATGCTTCTACAGATCTTACAGGAGAGGTTATTA 483
Db      498 CGGCCATGGGGCCCTTTCTGTACCATGCTTCTACAGATCTTACAGGAGAGGTTATTA 557

QY      484 GACCTCTGGACCTGCTTCCGGAGACTGCTGTAATCCGAGAAGACTGCCGGGGGAATATC 543
Db      558 GACCTCTGGACCTGCTTCCGGAGACTGCTGTAATCCGAGAAGACTGCCGGGGGAATATC 617

QY      544 CTGATTCGGGTCTCTCCAGAACCCATCAAGTATGCTTTGATTTTGGAGCGGCACTTC 603
Db      618 CTGATTCGGGTCTCTCCAGAACCCATCAAGTATGCTTTGATTTTGGAGCGGCACTTC 677

QY      604 CTGGCAGCAGTCCGAATCGGACTGTAGAGCCACCGGGCTCAACAGAGCGCTCTCTCCCGC 663
Db      678 CTGGCAGCAGTCCGAATCGGACTGTAGAGCCACCGGGCTCAACAGAGCGCTCTCTCCCGC 737

QY      664 AGTCATGCTGTGCTCTGCTCAAGGTGGACCGAGCGGGAAACGTTTGGGCCCCATTTGCGCAG 723

```

---

```

Db      738 AGTCATGCTGTGCTCTGCTGTCAGGTGGAACAGGGGGAACGTTTGGCCCCATTTGCGCAG 797
QY      724 CGAGAGGGAACAACTCTTACCTGATTGACTTGGCTGGGTGAGAGCAACCGCGGCACAGGC 783
Db      798 CGAGAGGGAACAACTCTTACCTGATTGACTTGGCTGGGTGAGAGCAACCGCGGCACAGGC 857
QY      784 AACAGGGCCCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGC 843
Db      858 AACAGGGCCCTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGC 917
QY      844 AAAGTGGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC 903
Db      918 AAAGTGGTAGATGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC 977
QY      904 ACTCGCCTATTGACAGGACTCTCTGGGTGGCTCAGCCACAGATATCTTATGCGCAACATT 963
Db      978 ACTCGCCTATTGACAGGACTCTCTGGGTGGCTCAGCCACAGATATCTTATGCGCAACATT 1037
QY      964 GCGCTGAGAGCGCTTCTACCTAGACACAGCTCTCCGACCTCAACTTTGCTGCCAGGTCC 1023
Db      1038 GCGCTGAGAGCGCTTCTACCTAGACACAGCTCTCCGACCTCAACTTTGCTGCCAGGTCC 1097
QY      1024 AAGGAGGTGATCAATCG 1040
Db      1098 AAGGAGGTGATCAATTTG 1114

RESULT 10
US-09-883-096-1
; Sequence 1, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid sequence of human kinesin motor
; OTHER INFORMATION: protein gene Hskip3a (Figure 1)
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a gene.
US-09-883-096-1

Query Match      9.5%; Score 138.6; DB 4; Length 4108;
Best Local Similarity 51.0%; Pred. No. 4.8e-30;
Matches 448; Conservative 0; Mismatches 404; Indels 27; Gaps 4;

QY      196 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGAGATGTCAGTTCAGTTCAGTTCAG 255
Db      344 TTTGACCGGTCTTTTGGCGAGGCGGCCACCCCAACAGGAGCTGTTCCAGCACACCCAGCAC 403
QY      256 CCCATCCTAAGCAGCTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTTCCCTATGACCCACA 315
Db      404 AGCGTCTCGGACAGCTTCTCTCCAGGCTACAACTGCTAGTGTGCTTACGGGGCCACC 463
QY      316 GGAAGCTGGGAAGACGACCAATGCTGGGAGCCAGAGCAACCTGGGTGTATCCCGCGG 375
Db      464 GGGGCTGGGAAGACACACACCATGCTGGGAAGGGAGGGGAGCCCGCGCATCATGTACCTG 523

```

Qy	376	GCTCTCATGGACCTCTCTCGAGCTCACAAGGAGGAGGGTGCCGAGGGCGCGCCATGGGCC	435
Db	524	AC-----CACCGTGGAACTGTACAGGGCCCTGGAGGCCCGCCAGCAGAGAACAC	574
Qy	436	CTTTCTGTCCACATGCTTTACTTAGAGATCTTACAGGAGAAGGTATTAGACCTCTTGAC	495
Db	575	TTCGAGGTGCTCATCAGCTACCCAGGAGGTGTATAATGAACAGATCCATGACCTCTCTGGAG	634
Qy	496	CCTGCTTCGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGGAATATCTCTCATTTCCGGGT	555
Db	635	C---C--AAAGGGGCCCTTGGCCATCCGCGAGGACCCCGCAAGGGGGTGGTGGTCAAGGA	691
Qy	556	CTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCACCTTCTGCCACGCCAGT	615
Db	692	CTTTCTTTCCACGACCGAGCCTCAGCCGAGCAGCTGCTGGAGATACTGACCAGGGGGAAC	751
Qy	616	CGAAATCGGACTGTAGGAGCCACCCGGCTCAAACAGCGCTCTCTCCCGCAGTCACTGCTGTG	675
Db	752	CGTAACCGCAGCGACACCCCACTGATGTCCAACGCACTTCTCTCCGCTCCCATGCCATC	811
Qy	676	CTCCCTGTCAAGGTGGACCGGGGAAGTTTGGCCCCCATTTTCCCGCAG-----CGAGAG	729
Db	812	TTCCAGATCTTTGTGAAGCAGCAGACCGGGTTCCAGGACTGACCCAGGCTGTCCAGGTG	871
Qy	730	GGAAACCTTACCTGATTTGACTTTGGCTGGGTTCAGAGGACAACCCGCGCAGACGGCAACAAG	789
Db	872	GCCAAGATGAGCTGATTTGACTTGGCTGGCTCAGAGCGGGCATCCAGCAGCCCATCGGAG	931
Qy	790	GGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTG	849
Db	932	GGGAGCGGCTCGGGAGGGGGCCAAATCAACCGCTCTCTGCTGGCGCTCATCAACGTC	991
Qy	850	GTAGATCGGCTG-----AATCAGGCGCTCCCTCGTGTACTTATTCGGAGCAGCAAG	900
Db	992	CTCAATCCCTTGGCCCGATGCAAAAGGGCGCAAGACCCATGTGTCCCTACCCGGGACAGCAA	1051
Qy	901	CTCACTCGCCTATTGCAAGGACTCTCTGGGTGGCTCAGCCACAGTATCCTTATTGGCCAAC	960
Db	1052	CTGACCGCCTGTCAAAGACTCCCTCGGGGGCAACTGCGCGCAAGTGTATGATCGCTGCC	1111
Qy	961	ATTGCCCCCTGAGAGACGCTTCTACTAGACAAGTCTCCGCACTTCAACTTTGCTGCCAGG	1020
Db	1112	ATCAGCCCCCTCCAGCCTGACCTACGAGGACACGTTACAACCCCTCAATATATGCCACCGG	1171
Qy	1021	TCCAGGAGGTGATCAATCGGCCTTTTACCAATGAGAGC	1059
Db	1172	GCCAAGAGATCAGGCTCTCGCTGAAGAGCAATGTGACC	1210

## RESULT 11

```

US-09-883-096-6
; Sequence 6, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a
;
; OTHER INFORMATION: fragment
; OTHER INFORMATION: Nucleotide sequence of HsKip3a fragment used in
; OTHER INFORMATION: ATPase assay (Figure 5).
US-09-883-096-6

```

Query Match	9.3%	Score 136.6;	DB 4;	Length 1152;
Best local Similarity	51.2%;	Pred. No. 9.5e-30;		
Matches 437;	Conservative 0;	Mismatches 389;	Indels 27;	Gaps 4;
QY	196	TTTGATGCGCTTCTATGGGAGAGAGGAGTACTCAGCAGGACATCTATGCAGGTTTCAGGTGCAG	255	
Db	214	TTTGA CCGGGTCTTTGGCGAGCGGGCCACCCAAACAGGACGTGTTCCAGACACACACGCAC	273	
QY	256	CCCATCCTAAGGCACTTGCTTGGAGGCGAGAAATGCCAGTGTGCTTGCTTATGCACCCACA	315	
Db	274	AGCGTCTCTGACAGCTTCCTCCAGGGCTCAACTGCTCTAGTGTTCCTACGGGGCCACC	333	
QY	316	GGAGCTGGGAGACGACACATGCTGGCAGGCCACAGAGCAACTCTGGGTGTATCCCGCG	375	
Db	334	GGGGCTGGGAAGACACACACCATGCTCTGGGAAGGAGGGGACCCCGGCATCATGTAT	389	
QY	376	GCTCTCATGACCTCTGTCAGAGCTCAACAAGGGAGGAGGTGCGAGGGCCGGCCATCGGCC	435	
Db	390	-----CTGACCACTGCTGNACTGTACAGCGCTCGAGGCCCGCCAGCAGGAGAAGCAC	444	
QY	436	CTTTCTGTCAACATGTTTACCTTAGAGATCTACAGGAGAAGGTATTAGACCTCTCTGGAC	495	
Db	445	TTCCAGGTGCTCATCAGCTACCAAGAGGTGTATATGAACAGATCCATGACCTCTCTGGAG	504	
QY	496	CTGCTTTGGGAGACCTGGTAAATCCGAGAAGACTCGCGGGGGAATATCTCTGATTCGGGT	555	
Db	505	C---CCAAAGGGGCCCTTGCCATCCCGAGGACCCCGCAAAAGGGGTGTGTGTCAGGA	561	
QY	556	CTCTCCAGAGGCCATCAGTAGCTTTTGCTGATTTTGTAGCGGCACCTTCCTGCCAGCCAGT	615	
Db	562	CTTTCTTTCCACAGCCAGCCTCAGCCGAGAGCTCTCTGAGATATCTGACCAGGGGGAAC	621	
QY	616	CGAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGCGCTCTCTCCGCGAGTCATGCTGTG	675	
Db	622	CGTAACCGCAGCAGCACCCCACTGATGCCAAACGGACCTTCCTCCCGCTCCCATGCCATC	681	
QY	676	CTCCTGGTCAAGGTGACAGCGGGAAAGTGTGGCCCCCATTTGCGCCAG-----CGAGAG	729	
Db	682	TTCCAGATCTTTGTGAAGCAGCAGGACCGGGTTCCAGGACTGACCCAGGCTGTCTCCAGGTG	741	
QY	730	GGAAACTCTACTGATTGACTTGGCTGGGTGAGAGGACAAACCGGGCGCAGGCAACAG	789	
Db	742	GCCAGATGAGCTGATTGACCTGGCTGGCTCAGAGCGGGCATCCAGACCCCATGCGAAG	801	
QY	790	GGCTTTCCGCTAAAGAGAGTGGAGCCATCAACACCTCCCTCTTTGCTCTGGCGAAAGTG	849	
Db	802	GGGAGCGCTCGGAGGGGGCCCAACATCAACCGCTCTCTCTGCTGGCGTCTCATACGTC	861	
QY	850	GTAGATGCGCTGAATCAGG-----GCCTCCCTCTGTGTACTTTATCGGAGACAAAG	900	
Db	862	CTCAATGCCCTTGGCCGATGCAAAAGGGCCGCAAGACCCATGTGCTCTACCGGACAGCAAA	921	
QY	901	CTCAGCTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTCTATTGCGAAC	960	
Db	922	CTGACCCGCTGCTCAAGACTCTCCCTCGGGGGCAACTGCGCCACAGTGTATGCTGCTGCC	981	
QY	961	ATTGCCCTTGAGAGCGCTTCTACCTTAGACACAGTCTCCGCACTCAACTTTGCTGCCAG	1020	
Db	982	ATCAGCCCCCTCAGCGCTGACCTACGAGGACAGTCAACACCCCTCAATATGCCGACCG	1041	
QY	1021	TCCAAGGAGGTGA	1033	
Db	1042	GCCAAGGAGATCA	1054	

RESULT 12  
US-09-621-233-1

[illegible]

Db 517 GGGAAACGGGAGAGAGACCCAGAGGCCAGCGCCGCAACCAAGAGAGTCTCCCGCTCCAC 576  
Qy 670 GCTGTGCTCTCTGCTCAAGGTGGACAGCGGAAAGTCTGGCCCATTTGCGCCAGAGAG 729  
Db 577 GCGGGCGCCCTGTTCA-----592  
Qy 730 GGAAGAACTCTACCTGATGACTTGGCTGGGTTCAGAGGACAAACCGCGCACAGGCAACAAG 789  
Db 593 -----TGATCGACCTGGCTGGCTCAGAGCGCGCTCGCAGACACAGAAATCGT 639  
Qy 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGAGT 849  
Db 640 GGGCAGCGTATGAAGAGGGGGCCACATCAACCGCTCACTGTGGCACTGGGCAACTGC 699  
Qy 850 GTAGATCGCTGAATCAG-----GGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903  
Db 700 ATCAACGCCCTGAGCGACAGGTAGCAACAGTATCACTATCGGACAGCAGCTC 759  
Qy 904 ACTCGCTATTCAGAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTTATTGGCAACATT 963  
Db 760 ACCGGCTCTGAAGGACTCTCTGGAGGAAACAGCGCACAGTATGATCGCTCACATC 819  
Qy 964 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1023  
Db 820 AGTCTCGAGCAGTGTGCTTTCGAGAGTCCCGAAACACCCCTGACCTACGCGCGCGGCC 879  
Qy 1024 AAGGAGTGTATCAATCGG 1041  
Db 880 AAGAACATTAAAGTAGG 897

## RESULT 14

US-09-724-516-1

; Sequence 1, Application US/09724516  
; Patent No. 6391573  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6391573el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1045  
; CURRENT APPLICATION NUMBER: US/09/724,516  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 09/621,233  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 897  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (472)...(495)  
; OTHER INFORMATION: n = a, c, t, or g  
US-09-724-516-1

Query Match 8.8%; Score 128.8; DB 3; Length 897;  
Best Local Similarity 50.1%; Pred. No. 1.6e-27;  
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 190 TACCAGTTTATGCTTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCGGTTCA 249  
Db 106 TACTGTTCAGCTGGCTTTGATCTTCCGCCACCCAGGAGATGGTGTATCAGGCCACC 165  
Qy 250 GTCCAGCCCATCTAAGGCACTTGTCTGGAAGGGCAGATCCAGTGTGCTTGCCTATGGA 309  
Db 166 ACCAAGAGCTCATCGAGGGCGTCACTCAGGCTACATGCCACTGTCTTTGCTATGGC 225  
Qy 310 CCACAGGAGCTGGGAGAGCGACACATGCTGGGAGCCCGACAGCAACTGGGGTGATC 369  
Db 226 CCCACAGGCTGTGGGAAACCTACACCATGCTGGGACAGACACAGGAGCGCTGGCATCTAT 285

Qy 370 CCGCGGCTCTCATGGACCTCTCGACGTCACRAAGGAGGAGGTGCGAGGCGCGGCCA 429  
Db 286 GTTCAGACCTCTCAAGACCTCTTCCGTGCCATCGAGGAGACCAATGA-----C 336  
Qy 430 TGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACCAGGAGAAAGGTATTAGACTC 489  
Db 337 ATGAGATGAGGTCTCCATGTCTACCTCGAGATCTACAATGAGATGATCCGGGACCTG 396  
Qy 490 CTGACCCCTCTTCGGGAGACCTGGTAATCGGAGAACTCGCGGGGGAATATCTCTGATT 549  
Db 397 CTGAACCCCTCCCTGGGCTACTTGGAGCTCGGGAGGACTCTAAGGGGGTGATCCAGGTG 456  
Qy 550 CCGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTTGTCTGATTTTGAAGCGCACTTCTGCA 609  
Db 457 GCGGCATCACCAANN 516  
Qy 610 GCCAGTCGAAATCGGACTGTAGAGGCCACCCGGCTCAACAGCGCTCCTCCCGCAGTCAT 669  
Db 517 GGGAAACCGGACAGGAGACCCAGAGGCCCAACCGCGCCCAACAGACGCTCTCTCCGCTCC 576  
Qy 670 GCTGTGCTCTCTGCTCAAGGTGGACCGGGAAGGTTTGGCCCCATTTGCGCAGCGAGAG 729  
Db 577 GCGGGCGCGCTGTCA-----592  
Qy 730 GGAAGAACTCTACCTGATTTGGCTGGGTTCAGAGGACAAACCGCGCACAGGCAACAAG 789  
Db 593 -----TGATCGACCTGGCTGGCTCAGAGCGCGCTCGCAGACACAGAAATCGT 639  
Qy 790 GGCCTTCGGCTAAAGAGAGTGGAGGCCATCAACACCTCTCTCTTGTCTGGGCAAGAGT 849  
Db 640 GGGCAGCGTATGAAGGAGGGGGCCACATCAACCGCTCACTGTGGCACTGGGCAACTGC 699  
Qy 850 GTAGATCGCTGAATCAG-----GGCTCCCTCGTGTACCTTATCGGAGACAGCAAGCTC 903  
Db 700 ATCAACGCCCTGAGCGACAGGAGGTAGCAACAGTATCACTATCGCAGACAGCAAGCTC 759  
Qy 904 ACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTTATTGGCAACATT 963  
Db 760 ACCGGCTCTGAAGGACTCTCTGGAGGAAACAGCGCACAGTATGATCGCTCACATC 819  
Qy 964 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACTCAACTTTGCTGCCAGTCC 1023  
Db 820 AGTCTCGAGCAGTGTGCTTTCGAGAGTCCCGAAACACCCCTGACCTACGCGCGCGGCC 879  
Qy 1024 AAGGAGTGTATCAATCGG 1041  
Db 880 AAGAACATTAAAGTAGG 897

## RESULT 15

US-10-090-695-1

; Sequence 1, Application US/10090695  
; Patent No. 6664072  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6664072el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1045  
; CURRENT APPLICATION NUMBER: US/10/090,695  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: 09/621,233  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 897  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (472)...(495)

; OTHER INFORMATION: n = a, c, t, or g  
US-10-090-695-1

```
Query Match      8.8%; Score 128.8; DB 4; Length 897;
Best Local Similarity 50.1%; Pred. No. 1.6e-27;
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

QY 190 TACAGTTTGTATGCTTCTATGGGAGAGAGTACTACAGCAGACATCTATCAGGTTCA 249
   |||||
Db 106 TACCTGTTGAGGTGGCTTTGACTTACCGCCACCCAGGAGATGTTGTATCAGGCCACC 165

QY 250 GTGACGCCATCCTAAGGCACCTTGTGGAAGGCGAGATGCCAGTGTGCTTGCCTATGGA 309
   |||||
Db 166 ACCAAGAGCCTCATCAGGGCGTCTCTCAGGCTACAATGCCACTGTCTTTGCCCTATGSC 225

QY 310 CCCACAGGAGCTGGGAGAGCAGACACAATGCTGGGCGAGCCACAGAGCAACCTGGGGTGATC 369
   |||||
Db 226 CCCACAGGCTGGGAGAGACCTACACCATGCTGGGCACAGACCAGGAGCCTGGCATCTAT 285

QY 370 CCGCGGGCTCTCATGACCTCTCTGACCTCAAGGGAGGAGGGTCCGAGGGCCGGCCA 429
   |||||
Db 286 GTTCAGACCCTCAACGACCTCTTCCGTGCCATCGAGGAGACCAGCAATGA-----C 336

QY 430 TGGGCCCTTCTGTCTACCATGCTTACTAGATCTACAGGAGAGGTATTTAGACCTC 489
   |||||
Db 337 ATGGAGTATGAGGTCTCCATGCTTACCTACCTGGAGATCTACAATGAGATGATCCGGGACCTG 396

QY 490 CTGGACCTCTCTCGGGAGACCTGGTAAATCCGAGAGAGACTGCCGGGGAAATATCCTGATT 549
   |||||
Db 397 CTGAAACCCCTCCCTGGGCTACCTGGAGCTCGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456

QY 550 CCGGGTCTCTCCAGAGGCCATCAGTAGTTTGTGATTTTGGAGGGGCACTTCTTGCCA 609
   |||||
Db 457 GCGGGCATACCCGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516

QY 610 GCCAGTCTGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCCCGCAGTCTAT 669
   |||||
Db 517 GGGAAACCGGAGAGGACCCAGGAGCCACCGGCCGCAACAGAGCTCTCCCGCTCCAC 576

QY 670 GCTGTGCTCTCTGATCAAGTGGACACAGCGGGAAACGTTTGGCCCCCATTTCCGCCAGGAG 729
   |||||
Db 577 GCGGGCGCGCTGTCA----- 592

QY 730 GGAAGACTCTACCTGATTGACTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 789
   |||||
Db 593 -----TGATCGACCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 639

QY 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAGTG 849
   |||||
Db 640 GGGCAGGCTATGAGAGGGGGGCCACATCAACCGCTCACTGTGGCAGTGGGCACTGC 699

QY 850 GTAGATCGCTGAATCAG-----GGCCTCCCTGTTGTTACCTTATCGGGACAGCAAGCTC 903
   |||||
Db 700 ATCAAGCCCTGAGCGCAAGGGTAGCAACAAGTACATCAACTATCGGACAGCAAGCTC 759

QY 904 ACTCGCTATGAGGACTCTCTGGTGGCTCAGCCACAGTATCTTATTGCCAATTT 963
   |||||
Db 760 ACCCGGCTCTGAGGACTCTCTGGGAGGAAACAGCGGCAAGTATGATCGCTTCATC 819

QY 964 GCCCTCAGAGACGCTTCTACCTAGACAGTCTCCGCACTCAACTTTTGTGCCAGGTCC 1023
   |||||
Db 820 AGTCTCGGAGAGTGGCTTCGAGGAGTCCCGGACACACCTTGACCTACGCGCGCGGCC 879

QY 1024 AAGAGGTGATCAATCGG 1041
   |||||
Db 880 AAGAACATTAAAGCTAGG 897
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 19:47:28 ; Search time 6469.64 Seconds  
(without alignments)  
10701.074 Million cell updates/sec

Title: US-10-797-893-7  
Perfect score: 1464  
Sequence: 1 atgggtcgctgcgctaag.....accattgtcccaaatgtga 1464

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464	100.0	1464	6 AR210055	AR210055 Sequence
2	1461	99.8	1538	6 AR210053	AR210053 Sequence
3	1458	99.6	1998	9 BT007259	BT007259 Homo sapi
4	1458	99.6	1998	12 BT007888	BT007888 Synthetic
5	1458	99.6	2097	6 AR304057	AR304057 Sequence
6	1458	99.6	2097	9 AB017430	AB017430 Homo sapi
7	1458	99.6	2117	9 BC028155	BC028155 Homo sapi
8	1458	99.6	2134	9 BC004352	BC004352 Homo sapi
9	1067.8	72.9	2086	10 BC003427	BC003427 Mus muscu
10	1038.4	70.9	1041	6 AR210054	AR210054 Sequence
11	1035.4	70.7	1115	6 AR210052	AR210052 Sequence
12	881.8	60.2	151041	2 AC101919	AC101919 Mus muscu
13	881.8	60.2	164759	10 AC101752	AC101752 Mus muscu
14	740.6	50.6	297639	2 AC101908	AC101908 Mus muscu
15	706.8	48.3	196674	10 AC102127	AC102127 Mus muscu
16	511.6	34.9	2756	5 BC063896	BC063896 Xenopus t
17	510.2	34.8	2583	5 XLA249841	XLA249841 Xenopus l
18	508.6	34.7	2565	5 BC043733	BC043733 Xenopus l
19	508.6	34.7	2640	5 BC073177	BC073177 Xenopus l

20	501	34.2	2593	5	AF267850	AF267850 Xenopus l
21	501	34.2	2661	5	XLA249840	XLA249840 Xenopus l
22	501	34.2	2781	5	BC070549	BC070549 Xenopus l
23	501	34.2	2819	5	AF267849	AF267849 Xenopus l
c 24	395	27.0	68916	2	AC101927	AC101927 Mus muscu
25	376.4	25.7	386	6	AX247007	AX247007 Sequence
c 26	365	24.9	174007	9	AC023831	AC023831 Homo sapi
c 27	365	24.9	174363	9	AC009133	AC009133 Homo sapi
28	361.8	24.7	3396	9	AB017333S2	AB017334 Homo sapi
29	353	24.1	105485	9	HUAC002301	AC002301 Homo sapi
30	346.4	23.7	444	10	AF013119	AF013119 Mus muscu
31	265	18.1	228783	2	AC094377	AC094377 Rattus no
c 32	265	18.1	235806	2	AC133764	AC133764 Rattus no
c 33	265	18.1	289504	2	AC123480	AC123480 Rattus no
c 34	262.4	17.9	195987	10	AC122537	AC122537 Mus muscu
c 35	262.4	17.9	226601	10	AC122863	AC122863 Mus muscu
36	208.8	14.3	1982	3	AK114179	AK114179 Ciona int
37	185.6	12.7	2095	6	AX833125	AX833125 Sequence
38	185.6	12.7	2095	9	AK094619	AK094619 Homo sapi
39	165	11.3	146415	5	AL929469	AL929469 Zebrafish
40	165	11.3	176906	2	CR354587	CR354587 Danio rer
41	153.8	10.5	3358	10	BC057614	BC057614 Mus muscu
42	145.8	10.0	435	5	MSU34658	U34658 Morone saxa
43	145.2	9.9	44004	2	AC092563	AC092563 Hyllobates
44	143	9.8	183586	9	AC092562	AC092562 Papio ham
45	139	9.5	9701	9	AB017333S3	AB017335 Homo sapi

ALIGNMENTS

RESULT 1  
AR210055  
LOCUS AR210055 1464 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 7 from patent US 6387644.  
ACCESSION AR210055  
VERSION AR210055.1 GI:21512186  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1464)  
AUTHORS Beraud,C.  
TITLE Motor proteins and methods for their use  
JOURNAL Patent: US 6387644-A 7 14-MAY-2002;  
FEATURES  
source Location/Qualifiers  
1..1464  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match	100.0%	Score 1464;	DB 6;	Length 1464;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1464;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGGTGCGTGTGCGGCTTAAGCAAGATTGGAGTACTCGTCTGCTCCACCTCCAGCTCGGTA	60	
Db	1	ATGGGTGCGTGTGCGGCTTAAGCAAGATTGGAGTACTCGTCTGCTCCACCTCCAGCTCGGTA	60	
Qy	61	AGGGTGGCTGTGCGACTCGGCCATTGTGGATGGAAACAGCGGGAGCAAGTGATCCCCCC	120	
Db	61	AGGGTGGCTGTGCGACTCGGCCATTGTGGATGGAAACAGCGGGAGCAAGTGATCCCCCC	120	
Qy	121	TGTGTGGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGAG	180	
Db	121	TGTGTGGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGAG	180	
Qy	181	ACTCTCAATACACAGTTTGATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT	240	
Db	181	ACTCTCAATACACAGTTTGATGCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTAT	240	
Qy	241	GCAGGTTTCAGTCAGCGCCCATCTTAAGCACTTCTGGAAGGGCAGATGCCAGTGTGCTT	300	

Db	241	GCAGGTTCA	GTCA	GTG	CGAGCCCAT	CTCTAA	GGCGCAT	CTTGCT	CGAAGGGCAG	ATATCC	CAGTGTGCTT	300		
Qy	301	GCCTAT	TGGACCC	CACAGAG	CTGGGA	AGAGCG	CACAA	TGCTGG	CGCACCC	CAGAGCA	ACT	360		
Db	301	GCCTAT	TGGACCC	CACAGAG	CTGGGA	AGAGCG	CACAA	TGCTGG	CGCACCC	CAGAGCA	ACT	360		
Qy	361	GGGGTG	ATATCC	CGGGGGT	CTCAT	TGGACCT	TCTG	CAGCT	CACA	GGGAGAGGGT	TCCGAG	420		
Db	361	GGGGTG	ATATCC	CGGGGGT	CTCAT	TGGACCT	TCTG	CAGCT	CACA	GGGAGAGGGT	TCCGAG	420		
Qy	421	GGCGGG	CCATGG	CGCCCTT	CTGT	CACCAT	GTCTTA	CTTAG	AGATCT	TAC	CAGGAGAGGTA	480		
Db	421	GGCGGG	CCATGG	CGCCCTT	CTGT	CACCAT	GTCTTA	CTTAG	AGATCT	TAC	CAGGAGAGGTA	480		
Qy	481	TTAGAC	CTCT	TGGACCT	TGCTT	CGGAGAC	CTTG	TAA	TCG	AGAA	CACTCC	CGGGGGA	540	
Db	481	TTAGAC	CTCT	TGGACCT	TGCTT	CGGAGAC	CTTG	TAA	TCG	AGAA	CACTCC	CGGGGGA	540	
Qy	541	ATCCTG	ATATCC	GGGTCT	CTCC	AGAGCC	CATCAG	TAGT	TTGCT	GTATTT	TGAGCGGC	600		
Db	541	ATCCTG	ATATCC	GGGTCT	CTCC	AGAGCC	CATCAG	TAGT	TTGCT	GTATTT	TGAGCGGC	600		
Qy	601	TTCTGT	CCAGC	AGTCG	AAATCG	CAAT	TGAG	AGCC	ACCG	GGTCTCA	AC	CAGCGCT	660	
Db	601	TTCTGT	CCAGC	AGTCG	AAATCG	CAAT	TGAG	AGCC	ACCG	GGTCTCA	AC	CAGCGCT	660	
Qy	661	CGCAGT	CA	TGCTG	CTCTG	GTCA	AGT	TGGA	CCG	GGGA	CGTT	TGGCCCCAT	720	
Db	661	CGCAGT	CA	TGCTG	CTCTG	GTCA	AGT	TGGA	CCG	GGGA	CGTT	TGGCCCCAT	720	
Qy	721	CAGCGA	GAGG	AAAACT	CTAC	CTG	AT	TGA	CTTG	GGCT	CGAG	GACAA	CGCGGC	780
Db	721	CAGCGA	GAGG	AAAACT	CTAC	CTG	AT	TGA	CTTG	GGCT	CGAG	GACAA	CGCGGC	780
Qy	781	GGCAAC	AA	GGGCC	TT	CGGCT	TAA	AGAG	TG	AGCC	ATCA	ACCT	CCCTG	840
Db	781	GGCAAC	AA	GGGCC	TT	CGGCT	TAA	AGAG	TG	AGCC	ATCA	ACCT	CCCTG	840
Qy	841	GGCAAG	TGGT	PAG	TGCG	CTGA	AT	CA	GGG	CTC	CGT	GTAC	TTAT	900
Db	841	GGCAAG	TGGT	PAG	TGCG	CTGA	AT	CA	GGG	CTC	CGT	GTAC	TTAT	900
Qy	901	CTCA	CT	CGCCT	TA	TTC	AGGA	CTCT	CGG	TGGCT	CAG	CC	CAC	960
Db	901	CTCA	CT	CGCCT	TA	TTC	AGGA	CTCT	CGG	TGGCT	CAG	CC	CAC	960
Qy	961	ATTGCC	CT	TGAG	AGCG	CTTCT	AC	TAG	AC	AGTCT	CG	CAC	TTTG	1020
Db	961	ATTGCC	CT	TGAG	AGCG	CTTCT	AC	TAG	AC	AGTCT	CG	CAC	TTTG	1020
Qy	1021	TCCA	AG	GAG	GTG	ATCA	T	CGG	CTTT	TAC	CAAT	GAG	AGCT	1080
Db	1021	TCCA	AG	GAG	GTG	ATCA	T	CGG	CTTT	TAC	CAAT	GAG	AGCT	1080
Qy	1081	CTGT	T	AA	GC	TG	CTC	A	AA	AG	AA	T	CG	1140
Db	1081	CTGT	T	AA	GC	TG	CTC	A	AA	AG	AA	T	CG	1140
Qy	1141	CCTG	AG	GA	GAG	GTG	GG	AG	CC	CTG	AG	CC	CT	1200
Db	1141	CCTG	AG	GA	GAG	GTG	GG	AG	CC	CTG	AG	CC	CT	1200
Qy	1201	AA	CT	C	AG	CC	CT	C	A	GA	AG	CT	T	1260
Db	1201	AA	CT	C	AG	CC	CT	C	A	GA	AG	CT	T	1260
Qy	1261	AG	CT	T	G	GA	CG	CT	T	G	CG	CT	C	1320
Db	1261	AG	CT	T	G	GA	CG	CT	T	G	CG	CT	C	1320
Qy	1321	A	AG	C	G	A	G	C	G	A	T	T	G	1380
Db	1321	A	AG	C	G	A	G	C	G	A	T	T	G	1380

Qy	1381	CTTAAAGCAGAAACAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTTGAGGAAAG	1440
Db	1381	CTTAAAGCAGAAACAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTTGAGGAAAG	1440
Qy	1441	GAGAACCATTTGTCCTCCACAATGTGA	1464
Db	1441	GAGAACCATTTGTCCTCCACAATGTGA	1464
RESULT 2			
LOCUS	AR210053	1538 bp	DNA
DEFINITION	Sequence 3 from patent US 6387644.		linear
ACCESSION	AR210053		
VERSION	AR210053.1	GI:21512183	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1538)		
TITLE	Beraud, C.		
JOURNAL	Motor proteins and methods for their use		
FEATURES	Patent: US 6387644-A 3 14-MAY-2002;		
source	Location/Qualifiers		
	1..1538		
	/organism="unknown"		
ORIGIN	/mol_type="unassigned DNA"		
Query Match	99.8%;	Score 1461;	DB 6; Length 1538;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1461;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	4	GGTGCTCTCGGCTAAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCGCGTAAGG	63
Db	78	GGTCGCTCTCGGCTAAAGCAAGATTGGAGTACTCGTCGTCCACCTCCAGCTCGCGTAAGG	137
Qy	64	GTGGCTGTGCACTCGCGCCATTTGTGATGGAAACAGCGGAGCAAGTGATCCCCCTGT	123
Db	138	GTGGCTGTGCACTCGCGCCATTTGTGATGGAAACAGCGGAGCAAGTGATCCCCCTGT	197
Qy	124	GTGCGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	183
Db	198	GTGCGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	257
Qy	184	CTCAAAATACCAAGTTTGTATGTCCTTTATGGGAGAGGAGTACTACAGCAGACAATCTATGCA	243
Db	258	CTCAAAATACCAAGTTTGTATGTCCTTTATGGGAGAGGAGTACTACAGCAGAGACTCTATGCA	317
Qy	244	GTTTCAGTGCAGCCCATCTTAAGGCACATCTGTTGAAAGGCGAGAAATGCCAGTGTGTTGCC	303
Db	318	GTTTCAGTGCAGCCCATCTTAAGGCACATCTGTTGAAAGGCGAGAAATGCCAGTGTGTTGCC	377
Qy	304	TATGGAACCCACAGGAGCTGGGAAGACGCACACAATGTCTGGGAGCCACAGAGCAACCTGGG	363
Db	378	TATGGAACCCACAGGAGCTGGGAAGACGCACACAATGTCTGGGAGCCACAGAGCAACCTGGG	437
Qy	364	GTGATCCCGCGGCTCTCATGAGCCTCTGTGAGCTCAACAGGAGAGGGTCCGAGGCG	423
Db	438	GTGATCCCGCGGCTCTCATGAGCCTCTGTGAGCTCAACAGGAGAGGGTCCGAGGCG	497
Qy	424	CGGCCATGGGCCCTTTCTGTACCATGCTTACCTAGAGATCTACAGAGAAAGGTATTATA	483
Db	498	CGGCCATGGGCCCTTTCTGTACCATGCTTACCTAGAGATCTACAGAGAAAGGTATTATA	557
Qy	484	GACCTCTCTGGACCTCTTCGGGAGACCTGGTAAATCCGAGAAAGACTGCCGGGGGAATATC	543
Db	558	GACCTCTCTGGACCTCTTCGGGAGACCTGGTAAATCCGAGAAAGACTGCCGGGGGAATATC	617
Qy	544	CTGATTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACCTTC	603
Db	618	CTGATTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACCTTC	677



QY	244	GGTTTCAGTTCAGCCATCTTAAGGCACTTGTCTGGAAGGCGCAATGCGCAGTGTGCTTCC	303
Db	316	GGTTTCAGTTCAGCCATCTTAAGGCACTTGTCTGGAAGGCGCAATGCGCAGTGTGCTTCC	375
QY	304	TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGCCACAGCAACCTGGG	363
Db	376	TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGCCACAGCAACCTGGG	435
QY	364	GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAGGAGGAGGGTCCGAGGCG	423
Db	436	GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAGGAGGAGGGTCCGAGGCG	495
QY	424	CGGCCATGGCGCCCTTCTGTCCACATGCTTACCTAGAGATCTACAGGAGAAGTATTA	483
Db	496	CGGCCATGGCGCCCTTCTGTCCACATGCTTACCTAGAGATCTACAGGAGAAGTATTA	555
QY	484	GACCTCTGGACCTCTCGGAGACCTGTAATCCGAGAGACTGCGCGGGGATATC	543
Db	556	GACCTCTGGACCTCTCGGAGACCTGTAATCCGAGAGACTGCGCGGGGATATC	615
QY	544	CTGATTCGGGCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGAGCGGCACTTC	603
Db	616	CTGATTCGGGCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGAGCGGCACTTC	675
QY	604	CTGCCAGCCAGTCCGAATCGGACTGTAGGAGCACCCCGGCTCAACAGCGCTCTCCCGC	663
Db	676	CTGCCAGCCAGTCCGAATCGGACTGTAGGAGCACCCCGGCTCAACAGCGCTCTCCCGC	735
QY	664	AGTCATGCTGTCTCTGTCAGGTTGGACCGGGAAGCTTTGGCCCCATTTGCCACAG	723
Db	736	AGTCATGCTGTCTCTGTCAGGTTGGACCGGGAAGCTTTGGCCCCATTTGCCACAG	795
QY	724	CGAGAGGGAATCTTACCTGATTGACTTGGCTGGGTTCAGAGGACCAACCGGCGCACAGC	783
Db	796	CGAGAGGGAATCTTACCTGATTGACTTGGCTGGGTTCAGAGGACCAACCGGCGCACAGC	855
QY	784	AACAAGGCGCTTCGGCTTAAAGAGAGTGGAGCATCAACCTCCCTGTTGCTTGGGC	843
Db	856	AACAAGGCGCTTCGGCTTAAAGAGAGTGGAGCATCAACCTCCCTGTTGCTTGGGC	915
QY	844	AAAGTGTAGATGGCTGTAATCAGGCGCTCCCTGCTGATCTTATCGGACACCAAGCTC	903
Db	916	AAAGTGTAGATGGCTGTAATCAGGCGCTCCCTGCTGATCTTATCGGACACCAAGCTC	975
QY	904	ACTCGCTTATGACGACTCTCTGGTGGCTCAGCCACAGTATCTTTATGTCACATTT	963
Db	976	ACTCGCTTATGACGACTCTCTGGTGGCTCAGCCACAGTATCTTTATGTCACATTT	1035
QY	964	CCCCCTGAGAGACCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGTCC	1023
Db	1036	CCCCCTGAGAGACCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGTCC	1095
QY	1024	AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCATGCTTGGGACCT	1083
Db	1096	AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCCTGACGCTCATGCTTGGGACCT	1155
QY	1084	GTTAAGTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGGCGCT	1143
Db	1156	GTTAAGTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGGCGCT	1215
QY	1144	GAGGAGGAGATTTGGAGCCCTGAGCCATGCGAGCTCCAGGCTCTGCTCCAGAAA	1203
Db	1216	GAGGAGGAGATTTGGAGCCCTGAGCCATGCGAGCTCCAGGCTCTGCTCCAGAAA	1275
QY	1204	CTCAGCCCTTACAGAACCTAAGCAGCATGAGCCCGGCTATGCTGGAGCGGCTCTCAGC	1263
Db	1276	CTCAGCCCTTACAGAACCTAAGCAGCATGAGCCCGGCTATGCTGGAGCGGCTCTCAGC	1335
QY	1264	TTGGAACCTGCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTAGTACCCCAAG	1323
Db	1336	TTGGAACCTGCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTAGTACCCCAAG	1395
QY	1324	CGAGAGCGGATGGTCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT	1383
Db	1396	CGAGAGCGGATGGTCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT	1455
QY	1384	AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGCCCCAGAGGCTGAGGAAAAGGAG	1443
Db	1456	AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGCCCCAGAGGCTGAGGAAAAGGAG	1515
QY	1444	AACCATTTGCCCAATG	1461
Db	1516	AACCATTTGCCCAATG	1533
RESULT 4			
BT007888		1998 bp mRNA linear SYN 13-MAY-2003	
LOCUS		Synthetic construct Homo sapiens kinesin-like 4 mRNA, partial cds.	
DEFINITION			
ACCESSION	BT007888		
VERSION	BT007888.1	GI:30584614	
KEYWORDS	FLI CDNA.		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 1998)		
AUTHORS	Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.		
TITLE	Cloning of human full-length CDSs in BD Creator(TM) System Donor vector		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1998)		
AUTHORS	Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA		
COMMENT	This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: http://bioinfo.clontech.com/orfclones.		
FEATURES	Location/Qualifiers		
source	1..1998		
	/organism="synthetic construct"		
	/mol_type="mRNA"		
	/db_xref="taxon:32630"		
	/clone="GH0024511.0"		
	/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"		
	/lab_host="DH5alpha Tl resistant"		
	/note="Vector: pDNR-Dual"		
CDS	1..>1998		
	/note="Mutations: 1997:Stop->Leu"		
	/codon_start=1		
	/transl_table=11		
	/product="Homo sapiens kinesin-like 4"		
	/protein_id="AAP36560.1"		
	/db_xref="GI:30584615"		
	/translation="NAAGSTQQRMRMAAASAAATSGAGRCRLSKIGATRRPPARV RVAVLRPFVDGTAGADPPCVGRMDSCLSEIANWRNHOTLKYQFDAYGRSTQOD IYAGSVQPIRLHLLLEGQNAVLAYGPTGAGKTHMLGSPQGVIPRALMDLLQLTRE EGAGRPWALSVMSTLEYIQEKYLLDLPASGDLVIREDCRGNLI1PGLSOKPTSSF ADPFRHFPASNRVTGATLRNQRSHAVLLVKVDQRLAPRQRGKLYLIDL A GSDNRTNGKGLRKESGINTSLFVLGKVDVALNQGLPRVPYRDSKRLRLQDSLG GSAHSLIANIAPRRFYLDITVSAFNARKEVINRPTNSLOPHALGPVKLSQKE LGPPPKAARGPEEBEIGPMPAPASQKLSPLQKLSMDPAMLRLLSLDRLL ASQSGOGAPLSTPKEKRVLMKTEVKLEIRLTKQKLEAKMLQAKBEKHC FIMLRPLSHRTVTGAKPLKXAVMPLQIQEQAASPNABIHLKNNKRRKRLSLDAL		

EPBEKAEBCWELQISPELLAHGROKILDLNLESGSARDLSLQRIQPKKAQILVGVREL  
HGPFQVEDLERVEGITGKQMESFLKANILGAAGORCGASL"

## ORIGIN

Query Match 99.6%; Score 1458; DB 12; Length 1998;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGGTAAGG	63
Db	76	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGGTAAGG	135
QY	64	GTGGCTGTGCGACTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGT	123
Db	136	GTGGCTGTGCGACTGCGGCCAATTTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGT	195
QY	124	GTGCGGGCATGGACAGCTGCTCTAGAGATTCTAACTCGAGGACCAACAGAGACT	183
Db	196	GTGCGGGCATGGACAGCTGCTCTAGAGATTCTAACTCGAGGACCAACAGAGACT	255
QY	184	CTCAATAACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
Db	256	CTCAATAACCAATTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	315
QY	244	GGTTGAGTGAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTCTGCC	303
Db	316	GGTTGAGTGAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTCTGCC	375
QY	304	TATGAGCCACAGAGCTGGGAAGCAGCACCAATGCTGGGAGGAGGGTCCGAGGGC	363
Db	376	TATGAGCCACAGAGCTGGGAAGCAGCACCAATGCTGGGAGGAGGGTCCGAGGGC	435
QY	364	GTGATCCCGGGCTCTCATGAGCTCTCTGAGCTCTCAAGGGAGGAGGGTCCGAGGGC	423
Db	436	GTGATCCCGGGCTCTCATGAGCTCTCTGAGCTCTCAAGGGAGGAGGGTCCGAGGGC	495
QY	424	CGGCGATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACAGGAGAAAGTATTA	483
Db	496	CGGCGATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACAGGAGAAAGTATTA	555
QY	484	GACCTCTGGACCTGCTTGGGAGACTGTAATCCGAGAGACTGCCGGGGAATATC	543
Db	556	GACCTCTGGACCTGCTTGGGAGACTGTAATCCGAGAGACTGCCGGGGAATATC	615
QY	544	CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC	603
Db	616	CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC	675
QY	604	CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	663
Db	676	CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	735
QY	664	AGTCATCTGTCTCTGCTCAAGTGGACCGAGCGGGAACGTTTGGGCCCATTTGCGCAG	723
Db	736	AGTCATCTGTCTCTGCTCAAGTGGACCGAGCGGGAACGTTTGGGCCCATTTGCGCAG	795
QY	724	CGAGAGGGAACCTCTACCTGATTGACTTGGCTCGGTCAGAGGACAAACCGCGCACAGGC	783
Db	796	CGAGAGGGAACCTCTACCTGATTGACTTGGCTCGGTCAGAGGACAAACCGCGCACAGGC	855
QY	784	AACAAGGGCTTCGGCTTAAAGAGAGTGGAGGCCATCAACACCTCTGTTGCTCGGGC	843
Db	856	AACAAGGGCTTCGGCTTAAAGAGAGTGGAGGCCATCAACACCTCTGTTGCTCGGGC	915
QY	844	AAAGTGTAGATGGCTGAATCAGGGCTCTGCTGTGATCTTATCGGAGCAGCAAGCTC	903
Db	916	AAAGTGTAGATGGCTGAATCAGGGCTCTGCTGTGATCTTATCGGAGCAGCAAGCTC	975
QY	904	ACTCGCTTATTTGAGGACTCTCTGGTGGCTCAGGCCACAGTAGTCTTATTTGCCAATT	963
Db	976	ACTCGCTTATTTGAGGACTCTCTGGTGGCTCAGGCCACAGTAGTCTTATTTGCCAATT	1035
QY	964	GCCCTTGAGAGCGCTTCTACCTAGACACAGTCTCTCCGCACTCAACTTTGCTGCCAGGTC	1023

Db	1036	GCCCTTGAGAGCGCTTCTACCTAGACACAGCTCTCGCACTCAACTTTGCTGCCAGGTC	1095
QY	1024	AAGAGGTGTCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGACCT	1083
Db	1096	AAGAGGTGTCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGACCT	1155
QY	1084	GTTAAGCTGTCTCAGAAAGAAATTCCTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT	1143
Db	1156	GTTAAGCTGTCTCAGAAAGAAATTCCTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT	1215
QY	1144	GAGAAAGAGAGATTTGGAGCCCTTGAGCCCATGTCAGCTCTCAGCTCTGCCCTCCAGAAA	1203
Db	1216	GAGAAAGAGAGATTTGGAGCCCTTGAGCCCATGTCAGCTCTGCCCTCTGCCCTCCAGAAA	1275
QY	1204	CTCAGCCCCCTACAGAGACTTAAGCAGCATGAGACCCGCGCATCTGAGAGCGCTCTCAGC	1263
Db	1276	CTCAGCCCCCTACAGAGACTTAAGCAGCATGAGACCCGCGCATCTGAGAGCGCTCTCAGC	1335
QY	1264	TTGACCGCTCTGCTTGGCTCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1323
Db	1336	TTGACCGCTCTGCTTGGCTCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1395
QY	1324	CGAGCGGATGCTGCTTAATGAAGCAGTAGTAGAAGAGGACCTTAGAGATTGAGAGGCTT	1383
Db	1396	CGAGCGGATGCTGCTTAATGAAGCAGTAGTAGAAGAGGACCTTAGAGATTGAGAGGCTT	1455
QY	1384	AAGACGAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGAGAAAGAG	1443
Db	1456	AAGACGAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGGCTGAGAGAAAGAG	1515
QY	1444	AACATTGTCCCAATG 1461	
Db	1516	AACATTGTCCCAATG 1533	

## RESULT 5

AR304057

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1. 2097

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Query Match

Best Local Similarity

Matches 1458;

QY

4

GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGGTAAGG

63

Db

100

GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGGTAAGG

159

QY

64

GTGGCTGTGCGACTGCGGCCAATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCCCTGT

123

Db

160

GTGGCTGTGCGACTGCGGCCAATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCCCTGT

219

QY

124

GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAACACACAGGAGACT

183

Db

220

GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAACACACAGGAGACT

279

AR304057 2097 bp DNA linear PAT 12-JUN-2003  
Sequence 34 from patent US 6544766.

AR304057  
AR304057.1 GI:31692955  
Unknown.  
Unknown.  
Unclassified.

1 (bases 1 to 2097)  
Beraud, C., Ohashi, C., Sakowicz, R., Vaissberg, E., Wood, K. and Yu, M.  
Human kinesins and methods of producing and purifying human  
kinesins

Patent: US 6544766-A 34 08-APR-2003;  
Location/Qualifiers

1. 2097  
/organism="unknown"  
/mol\_type="genomic DNA"

Query Match 99.6%; Score 1458; DB 6; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGGTAAGG 63

Db 100 GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGGTAAGG 159

QY 64 GTGGCTGTGCGACTGCGGCCAATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCCCTGT 123

Db 160 GTGGCTGTGCGACTGCGGCCAATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCCCTGT 219

QY 124 GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAACACACAGGAGACT 183

Db 220 GTGCGGGCATGACAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAACACACAGGAGACT 279

Db 1360 TTGGACCGTCTGCTTGGCTCCAGGGGAGCCAGGGGGCCCCCTCTGTGTGAGTACCCCAAAG 1411

QY 1324 CGAGACCGATGGTGTCTAATGAGACAGTAGAGAGAGAGAGACCTAGAGATTGAGAGGCTT 1383

Db 1420 CGAGACCGATGGTGTCTAATGAGACAGTAGAGAGAGAGACCTAGAGATTGAGAGGCTT 1479

QY 1384 AGAGACCAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCCAGAGGCTGAGGAAAAGGAG 1443

Db 1480 AGAGACGAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCCAGAGGCTGAGGAAAAGGAG 1539

QY 1444 AACCATTTGCCCCAATG 1461

Db 1540 AACCATTTGCCCCAATG 1557

RESULT 6

AB017430

LOCUS 2097 bp mRNA linear PRI 06-MAR-1999

DEFINITION Homo sapiens mRNA for kinesin-like DNA binding protein, complete cds.

ACCESSION AB017430 D38751

VERSION AB017430.2 GI:4519442

KEYWORDS Kid; kinesin-like DNA binding protein; kinesin family.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (sites) Tokai,N., Fujimoto-Nishiyama,A., Toyoshima,Y., Yonemura,S., Tsukita,S., Inoue,Y. and Yamamoto,A.I.

TITLE Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and the mitotic spindle

JOURNAL EMBO J. 15 (3), 457-467 (1996)

MEDLINE 96174806

PUBMED 8599929

REFERENCE 2 (bases 1 to 2097)

AUTHORS Tokai-Nishizumi,N. and Edamasu,M.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-1998) Noriko Tokai-Nishizumi, The University of Tokyo, The Institute of Medical Science, Department of Oncology; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

COMMENT (E-mail:tokai@ngc.ims.u-tokyo.ac.jp, Tel:03-5449-5302, Fax:03-5449-5413)

FEATURES

Sequence updated (02-Nov-1994)

Location/Qualifiers

1..2097

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="pBSkid"

/cell\_line="MDA-MB453"

/cell\_type="breast cancer"

1..2097

/gene="Kid"

25..2022

/gene="Kid"

/codon\_start=1

/product="kinesin-like DNA binding protein"

/protein\_id="BAA33019.2"

/db\_xref="GI:4519443"

/translation="MAAGGTCQRRREMAAASAAASGACRLSKIGATRRPPPAR  
RVAVRLPFVDTGASDPCCVRGMDSCSLEIANWRNQETLKQFDAPYGERSTCO  
IYAGSVQPIRLHLEGGQNASVLAYGPTGAKTHWMLGSPQGVIPRALMDQLQTR  
EAGQRPWALSVTYSLEYIQEKVLDDPASGDIVREDCRGNILIPGLSKPILSL  
.ADFERHFTPASNRVTGASATRLNQRSRSHAVLLVKVDQERLAPFRQREKLYLIDL  
GSEDNRFTNGKRLKESGATISLFLVGLVDALNOQLRPVPRYRDSKTRLLQDLSL  
GSAISILANTAPERFYLDTVSALNFAARKEVINRPTNESLOPHALGPVKLSQK  
LLGPPEAKRGPEEERIGSPFPMAPAAK  
ASQSQGAPLLSTPKPRERNVLMTVEEKDLELUKTKQKELEAKMLQAKABEKNH  
PTMLRPLSHRTVTPKAPKLVKAAVVMPLQITQEOAASNAEIHILKNKGRKRLKSLDNL  
EPEKABDCWELQISPELLAHGRQKILLNNGSARDLSLQRI GPKKAQLIVGMRE

gene

CDS

HGPFSQVEDLERVEGITGKQMESFLKANILGLAAGQRCGAS"

```
Query Match          99.6%; Score 1458; DB 9; Length 2097;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	4	GGT	CGCTGT	CGGCTAAG	CAAGATTGG	AGACTACT	CGTCTGCTCA	CTTCCAGCTCGCGTAAGG	63
Db	100	GGT	CGCTGT	CGGCTAAG	CAAGATTGG	AGCTACT	CGTCTGCTCA	CTTCCAGCTCGCGTAAGG	159
Qy	64	GTG	GCTGTG	CACTCGG	CCATTGTG	TGATGGAA	CAGCGGAGCA	AGTGATGCCCCCTGT	123
Db	160	GTG	GCTGTG	CACTCGG	CCATTGTG	TGATGGAA	CAGCGGAGCA	AGTGATGCCCCCTGT	219
Qy	124	GTG	CGGGCAT	TGGA	CAGCTGCTCT	TAGAGATTGCTAA	CTGGAGGAAC	CAACGAGAGACT	183
Db	220	GTG	CGGGCAT	TGGA	CAGCTGCTCT	TAGAGATTGCTAA	CTGGAGGAAC	CAACGAGAGACT	279
Qy	184	CTCAAA	TACCACTTT	TGATGCTCT	TATGGGAGAGG	AGTAGTCTAC	GACGGACAT	CTATGCA	243
Db	280	CTCAAA	TACCACTTT	TGATGCTCT	TATGGGAGAGG	AGTAGTCTAC	GACGGACAT	CTATGCA	339
Qy	244	GGT	TCAGTGCA	CCCATCTTA	AGGCATCTTGCTGGA	AGGCAGAAATG	CCAGTGTGCTTGCC	303	
Db	340	GGT	TCAGTGCA	CCCATCTTA	AGGCATCTTGCTGGA	AGGCAGAAATG	CCAGTGTGCTTGCC	399	
Qy	304	TATGGA	CCCA	CAGAGCTGG	GAAGACGACACA	ATATGCTGGG	CAGCCACGAGCA	ACTTGGG	363
Db	400	TATGGA	CCCA	CAGAGCTGG	GAAGACGACACA	ATATGCTGGG	CAGCCACGAGCA	ACTTGGG	459
Qy	364	GTGAT	CTCCGGGGCTCT	CATGGA	CCTCTCTG	CAGCTCTCA	CAAGGGAGG	GGTCCGAGGGC	423
Db	460	GTGAT	CTCCGGGGCTCT	CATGGA	CCTCTCTG	CAGCTCTCA	CAAGGGAGG	GGTCCGAGGGC	519
Qy	424	CGG	CCATGGG	CCCTTTCTGT	CACCATGTCTT	ACTAGAGATCTT	ACCAAGGAAGT	ATTA	483
Db	520	CGG	CCATGGG	CCCTTTCTGT	CACCATGTCTT	ACTAGAGATCTT	ACCAAGGAAGT	ATTA	579
Qy	484	GAC	CTCTGGA	CCCTGCTT	CGGGAGACCTG	TGTAATCGAG	AAAGACTG	CCGGGGAAATATC	543
Db	580	GAC	CTCTGGA	CCCTGCTT	CGGGAGACCTG	TGTAATCGAG	AAAGACTG	CCGGGGAAATATC	639
Qy	544	CTGAT	TCGGGTCTCT	CCCCA	AGCCCATCAG	TAGTCTTGCTGATTTG	TGAGCGGCA	CTTC	603
Db	640	CTGAT	TCGGGTCTCT	CCCCA	AGCCCATCAG	TAGTCTTGCTGATTTG	TGAGCGGCA	CTTC	699
Qy	604	CTG	CCAGCCAGT	CGAAATCG	AGACTGTAG	GAGCCACCCGGCTCA	ACGAGCGCTCT	CCCGC	663
Db	700	CTG	CCAGCCAGT	CGAAATCG	AGACTGTAG	GAGCCACCCGGCTCA	ACGAGCGCTCT	CCCGC	759
Qy	664	AGT	CATGCTGTG	CTCTGGTCA	AGGTGGAC	CAGCGGAA	CGTTTGGCCCA	ATTTGCGCAG	723
Db	760	AGT	CATGCTGTG	CTCTGGTCA	AGGTGGAC	CAGCGGAA	CGTTTGGCCCA	ATTTGCGCAG	819
Qy	724	CGA	GAGGGA	AACTCTAC	CTGATTGACTT	GGCTGGGT	CAGAGGCA	CAACCGCGCAGGC	783
Db	820	CGA	GAGGGA	AACTCTAC	CTGATTGACTT	GGCTGGGT	CAGAGGCA	CAACCGCGCAGGC	879
Qy	784	AACA	GGGCTCTCGGCT	AAAA	GAGATGG	GAGCCATCA	ACACCTCCCTGTTGT	TGCTCGGGC	843
Db	880	AACA	GGGCTCTCGGCT	AAAA	GAGATGG	GAGCCATCA	ACACCTCCCTGTTGT	TGCTCGGGC	939
Qy	844	AAAG	TGTAGATGG	CGTGAAT	CAGGG	CGCTCCCTGCTGTAT	CGGTATCGGG	GACGAA	903
Db	940	AAAG	TGTAGATGG	CGTGAAT	CAGGG	CGCTCCCTGCTGTAT	CGGTATCGGG	GACGAA	999
Qy	904	ACT	CGCCTATTG	CAGGACTCTCT	GGGTGGCT	CTAGCC	CAAGTATCTT	ATTGGCCAACT	963
Db	1000	ACT	CGCCTATTG	CAGGACTCTCT	GGGTGGCT	CTAGCC	CAAGTATCTT	ATTGGCCAACT	1059
Qy	964	GCCC	CTGAGAGA	CGCTTCTACT	TAGACAC	AGTCTCG	CACTCA	ACTTGTGCGCAGGTC	1023

Db	1060	GC	CCCTGAGAGACGCTTTCTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAGGTCC	1119
Qy	1024	AAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGCAGCCTCATGCTTTGGGACCT	1083	
Db	1120	AAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCCTGCAGCCTCATGCTTTGGGACCT	1179	
Qy	1084	GTTTAAGCTGTCTCAGAAAGAAATTGCTTTGTTGTCCAACGAGGCCAAAGAGAGCCCGAGGCGCT	1143	
Db	1180	GTTTAAGCTGTCTCAGAAAGAAATTGCTTTGTTGTCCAACGAGGCCAAAGAGAGCCCGAGGCGCT	1239	
Qy	1144	GAGGAAGAGAGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCTGCTCTCCAGAAA	1203	
Db	1240	GAGGAAGAGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCTGCTCTCCAGAAA	1299	
Qy	1204	CTCAGCCCCCTCAGAAAGCTTAAGCAGATGGACCCGGGCCATGTGGAGCGCCTCTCTCAGC	1263	
Db	1300	CTCAGCCCCCTCAGAAAGCTTAAGCAGATGGACCCGGGCCATGTGGAGCGCCTCTCTCAGC	1359	
Qy	1264	TTGCAACGTCGTCTGCTCTCCAGGGGAGCCAGGGGGCCCCCTCTGTTTGAGTACCCCAAAG	1323	
Db	1360	TTGCAACGTCGTCTGCTCTCCAGGGGAGCCAGGGGGCCCCCTCTGTTTGAGTACCCCAAAG	1419	
Qy	1324	CGAGACGGATGTGCTTAATGAAGACAGTAGAAGAAAGACCTAGAGATTGAGAGCGCTT	1383	
Db	1420	CGAGACGGATGTGCTTAATGAAGACAGTAGAAGAAAGACCTAGAGATTGAGAGCGCTT	1479	
Qy	1384	AAGACCAAGCAAAAGAACTGGAGGCCAGATGCTTCGCCCAAGAGCTGAGGAAGAGGAG	1443	
Db	1480	AAGACCAAGCAAAAGAACTGGAGGCCAGATGCTTCGCCCAAGAGCTGAGGAAGAGGAG	1539	
Qy	1444	AACCATTTGTCCTCAATG	1461	
Db	1540	AACCATTTGTCCTCAATG	1557	

RESULT 7	BC028155	2117 bp	mRNA	linear	PRI 29-JUN-2004
LOCUS	BC028155				
DEFINITION	Homo sapiens kinesin family member 22, mRNA (cDNA clone MGC:40049 IMAGE:5241557), complete cds.				
ACCESSION	BC028155				
VERSION	BC028155.1	GI:20380446			
KEYWORDS	MGC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2117)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusan,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Rata,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	1247932				
REFERENCE	2 (bases 1 to 2117)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				



JOURNAL	Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabbs@mail.nih.gov">cgabbs@mail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc.mgc@nih.gov">nisc.mgc@nih.gov</a> Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Manduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.		
FEATURES	Clon distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 62 Row: 1 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein. Location/Qualifiers 1. .2117 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:40049 IMAGE:5241557" /tissue_type="Brain, fetal, whole pooled" /clone_lib="NIH MGC_121" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6" 1. .2117 /gene="KIF22" /notes="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4" /db_xref="LocusID:3835" /db_xref="MIM:603213" 6. .2003 /gene="KIF22" /codon_start=1 /product="kinesin family member 22" /protein_id="AAH28155.1" /db_xref="GI:20380447" /db_xref="LocusID:3835" /db_xref="MIM:603213" /translation="MAAGGTQQRREMAAASAAISGACRCLSKIGATRRPPPARV RVAVLPFPVDTAGADPPVGRMDSCLEIANRNHOETLVQYDFADYGERSTQD IYAGVOPILRHLEGQNSVLAYPEGTGAKTHTMLGSPEQGVIPRALMDLQLQTR EAGERWALSVMSYDEIYOEKVLDDLPASGLDVIREDRCNILLPGLSQPISSF ADPRFLPAGNRVTGATRLNQRSSHAVLLVVKVDQRLAPFRQREKQLIDLA GSDNRRTPKNGRLKESGAINTSFLVGLKVDLNLQGLPRVPRDSKLTLLQDSLG GSASHILIANIAPRRFYLDVTSALNPAARKEVINRPTFNESLQPHALGPVKLSQKE LLGPPEAKRARGPEEIEIGSPPEMAAPASOKLSPLQKLSMDPAMLERLLSLDRLL ASQSQCAPLLSTPKRPMVIMKTVEEKOLEIRLTKQKELEAKMLAOKAEKENHC PTMLRPLSHRTVQAKLKKAVVMPQLIOEQASPNAEIHLKNKRKEKRLSLDAL EPEBAEDCWELOISPELLAHGRQKILLOLNEGARDLRSIQRIQPKKAQLIVGREL HGPFPSQVEDLERVEGITQKQMESFLKANILGLAAGQRCGAS"		
gene	1. .2117 /gene="KIF22" /notes="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4" /db_xref="LocusID:3835" /db_xref="MIM:603213"		
CDS	6. .2003 /gene="KIF22" /codon_start=1 /product="kinesin family member 22" /protein_id="AAH28155.1" /db_xref="GI:20380447" /db_xref="LocusID:3835" /db_xref="MIM:603213"		
ORIGIN	Query Match 99.6%; Score 1458; DB 9; Length 2117; Best Local Similarity 100.0%; Pred. No. 0; Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 4 GGTCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCAGCTCCGCTAAGG 63 		
	Db	Qy	81 GGTCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCAGCTCCGCTAAGG 140
	Qy	Qy	64 GTGCTGTGCGACTGCGGCGCATTTGTGATGGAACAGCGGAGCAAGTGTATCCGCTGT 123
	Db	Db	141 GTGCTGTGCGACTGCGGCGCATTTGTGATGGAACAGCGGAGCAAGTGTATCCGCTGT. 200
	Qy	Qy	124 GTGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAACGAGACT 183
	Db	Db	201 GTGCGGGCATGGAGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAACGAGACT 260
	Qy	Qy	184 CTCAATACCACTTTGATGCTCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
	Db	Db	261 CTCAATACCACTTTGATGCTCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 320
	Qy	Qy	244 GGTTCAGTGCAGCCATCTTAAGCCTTGTGGAAGGCGAGATGCCAGTGTCTTGCC 303
	Db	Db	321 GGTTCAGTGCAGCCATCTTAAGCCTTGTGGAAGGCGAGATGCCAGTGTCTTGCC 380
	Qy	Qy	304 TATGAGCCCAAGAGCTGGGAGACGACACATATGCTGGGCGAGCCAGAGCAACCTGGG 363
	Db	Db	381 TATGAGCCCAAGAGCTGGGAGACGACACATATGCTGGGCGAGCCAGAGCAACCTGGG 440
	Qy	Qy	364 GTGATCCCGGGCTCTCATGACCTCTGAGCTCAAGGAGGAGGTCGCGAGGC 423
	Db	Db	441 GTGATCCCGGGCTCTCATGACCTCTGAGCTCAAGGAGGAGGTCGCGAGGC 500
	Qy	Qy	424 CGGCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 483
	Db	Db	501 CGGCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAAAGTATTA 560
	Qy	Qy	484 GACCTCTCGGACCTGCTTCGGGAGACCTGGTAACTCGAGAGAGTCCGCGGGGAATATC 543
	Db	Db	561 GACCTCTCGGACCTGCTTCGGGAGACCTGGTAACTCGAGAGAGTCCGCGGGGAATATC 620
	Qy	Qy	544 GTGATTCGGGTCTCTCCAGAGCCCATAGTAGCTTGTGATTTTGGCGGCACATTC 603
	Db	Db	621 GTGATTCGGGTCTCTCCAGAGCCCATAGTAGCTTGTGATTTTGGCGGCACATTC 680
	Qy	Qy	604 GTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663
	Db	Db	681 GTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 740
	Qy	Qy	664 AGTCATGTGTGCTCTCTGTCAGGTGGAACAGCGGGAACGTTTGGCCCCATTTGCCAG 723
	Db	Db	741 AGTCATGTGTGCTCTCTGTCAGGTGGAACAGCGGGAACGTTTGGCCCCATTTGCCAG 800
	Qy	Qy	724 CGAGAGGGAACCTCTACTGATTCAGCTTGGTGGTTCAGAGCAACCGGCGCACAGGC 783
	Db	Db	801 CGAGAGGGAACCTCTACTGATTCAGCTTGGTGGTTCAGAGCAACCGGCGCACAGGC 860
	Qy	Qy	784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTGTTGCTCTGGGC 843
	Db	Db	861 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTGTTGCTCTGGGC 920
	Qy	Qy	844 AAAGTGTAGATCGCTGATCAGGGCTCTCTCTGTCGTCCTTATCGGAGACGAGCTC 903
	Db	Db	921 AAAGTGTAGATCGCTGATCAGGGCTCTCTCTGTCGTCCTTATCGGAGACGAGCTC 980
	Qy	Qy	904 ACTCGCTATTCGAGACTCTCTGCTGGTGGCTCAGCGCCACAGTATCTCTATTGCGCAACT 963
	Db	Db	981 ACTCGCTATTCGAGACTCTCTGCTGGTGGCTCAGCGCCACAGTATCTCTATTGCGCAACT 1040
	Qy	Qy	964 GCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTCTGCTGCGAGTCC 1023
	Db	Db	1041 GCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTCTGCTGCGAGTCC 1100
	Qy	Qy	1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTCTGGAGCTT 1083
	Db	Db	1101 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTCTGGAGCTT 1160
	Qy	Qy	1084 GTTAAGCTGTCTCAGAAAGAAATTCGTTGGTCCACAGAGGCAAGAGAGCCGAGGCCCT 1143
	Db	Db	1161 GTTAAGCTGTCTCAGAAAGAAATTCGTTGGTCCACAGAGGCAAGAGAGCCGAGGCCCT 1220



1144 GAGGAAGAGGAGATTGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCTGCTCCAGAAA 1203  
 1221 GAGGAAGAGGAGATTGGAGCCCTGAGCCCATGGCAGCTCCAGCCTCTGCTCCAGAAA 1280  
 1204 CTGAGCCCTTACAGAGCTAAGCAGATGAGACCGCCGATCTGAGAGCGCTCTCTCAGC 1263  
 1281 CTCAGCCCTTACAGAGCTAAGCAGATGAGACCGCCGATCTGAGAGCGCTCTCTCAGC 1340  
 1264 TTGACCGCTCTGCTTGGCTCCAGCGAGCCAGCGGCGCCCTCTGTTGAGTACCCCAAG 1323  
 1341 TTGACCGCTCTGCTTGGCTCCAGCGAGCCAGCGGCGCCCTCTGTTGAGTACCCCAAG 1400  
 1324 CGAGCGGATGGTCTTAATGAAGACAGTAGAAGAGGACCTTAGAGATTGAGAGCCTT 1383  
 1401 CGAGCGGATGGTCTTAATGAAGACAGTAGAAGAGGACCTTAGAGATTGAGAGCCTT 1460  
 1384 AAGCAGGACCAAGAAAGTGGAGCCCAAGATGTTGCCCAAGAGCTGAGGAAAAGGAG 1443  
 1461 AAGCAGGACCAAGAAAGTGGAGCCCAAGATGTTGCCCAAGAGCTGAGGAAAAGGAG 1520  
 1444 AACCATGTGCCCAATG 1461  
 1521 AACCATGTGCCCAATG 1538

BC004352 2134 bp mRNA linear PRI 29-JUN-2004  
 IMAGS:3535435, complete cds.

ACCESSION BC004352  
 VERSION BC004352.1 GI:13279307  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2134)  
 Klausner,R.D., Collins,F.S., Wagner,L.H., Derge,J.G., Altschul,S.F., Zeeberg,B., Buettow,K.H., Shennan,C.M., Schuler,G.D., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McSwan,P.J., McKernan,K.J., Malek,D.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Young,A.C., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Greenchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Shreenko,Y., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 2134)  
 Strausberg,R.  
 Direct Submission  
 Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgi.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [gcapsb-remail.nih.gov](mailto:gcapsb-remail.nih.gov)  
 Tissue Procurement: DCTD/DTP  
 cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcsc.bc.ca  
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
 Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth  
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,  
 Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,  
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhhu,  
 Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska,  
 Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquele  
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 8 Row: f Column: 1  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 6453817.

FEATURES  
 Location/Qualifiers  
 1..2134  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:1573 IMAGE:3535435"  
 /tissue type="Lung, small cell carcinoma"  
 /clone lib="NIH MGC 7"  
 /lab host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..2134  
 /genes="KIF22"  
 /notes="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"  
 /db\_xref="LocusID:3835"  
 /db\_xref="MIM:603213"  
 20..2017  
 /genes="KIF22"  
 /codon\_start=1  
 /product="kinesin family member 22"  
 /protein\_id="AAH04352.1"  
 /db\_xref="GI:13279308"  
 /db\_xref="LocusID:3835"  
 /db\_xref="MIM:603213"  
 /translation="MAAGGSTQORREMAAASAAASGACRCLSKIGATRRPPPARV  
 RVAVRLRPFDGTAGADPPCVGRMDSCSLTANWNRHOBTKYQFDALFYGERSTQOD  
 IVAGSVQPTLRHLLEGONASVLAYGTGAGTKTLMGSPQEPQVIFRALMDLQLTR  
 EGAGRPWALSVTMSVLEIYQEKVLDLPASGDVIREDCRGNILPGLSQPISF  
 ADPERHFLPASNRNTRVATRLNQRSSRAVLVQDQERLAPFRQREKTLFLDLA  
 GSEDNRNKGKLGKSAINTSLFVKGKVDALNQGPRVPRVDSKLTRLQDSLG  
 GSAHSLILIANIAPERFRFYLDTSALNFAARKEVINRPTNESLQPHALGPVLSQKE  
 LIGPPEAKRGPEEEIGSPPEPMAAPASAKLSPLQSLSDMPALRLLSLDRLL  
 ASQSGCAPLLSTPKRERMVMTVEEKOLEIRLTKOKELAKLAOKAEKENHC  
 PTMLRPLSHRTVTGAKPLKAVVMPQLIQOQASPAETHILKNKGRKRLSLDAL  
 EBEKADCDWELQISPELLAHGRQKLDLNLNGBSARDLSLRIGPKAKQLIYGVREL  
 HGFPSQVEDLERVEGITGKMESFLKANILGLAAGQRCGAS"

Query Match 99.6%; Score 1458; DB 9; Length 2134;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGTGCGTGTGCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 63  
 95 GGTGCGTGTGCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 154  
 QY 64 GTGCGTGTGCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 123  
 155 GTGCGTGTGCGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGG 214  
 QY 124 GTGCGGGGATGACAGAGCTGCTCTAGAGATTGCTAAGTGGAGAACCAACAGGAGACT 183  
 215 GTGCGGGGATGACAGAGCTGCTCTAGAGATTGCTAAGTGGAGAACCAACAGGAGACT 274  
 QY 184 CTCAATACCAAGTTTGCATGCCTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243



Series: IRAK Plate: 5 Row: m Column: 7  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704181.

# FEATURES

Location/Qualifiers  
1..2086  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clones="MGC:6456 IMAGE:2615715"  
/tissue type="mammary tumor, C3(1)-Tag model, infiltrating ductal carcinoma 5 month old virgin mouse."  
/clone\_lib="NCI CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
1..2086  
/gene="Kif22"  
/notes="synonym: Kid"  
/db\_xref="LocusID:110033"  
/db\_xref="MGI:109233"  
20..2002  
/gene="Kif22"  
/codon\_start=1  
/product="kinesin family member 22"  
/protein\_id="AAH03427.1"  
/db\_xref="GI:13097360"  
/db\_xref="LocusID:110033"  
/db\_xref="MGI:109233"  
/translation="MSTRAKTCORREMASATSPGRCVSGKGLRPPPLARVRAVR  
LRPMDETEAKELPCVRAIDSCSLEVANKKQOBTUKYQDAFIGSKSQEYVGS  
VQILRLHLLGQNSVLAIGTGTAGKTHLMGSPQGVPIRALMDLLQALRESABG  
RPMDSVSYALEYIQKVLDDLPASGDLVIREDCRNILIPGLTKPITSFSDPQ  
HFLPASNRVAGTLMNQRSRSHVLLVKVDORERLTPRQREGLYLIDLAGSEB  
RRTGNQIRLKESGAINISFLVGLVDALNOGLPIPYRDSKLTRILQSLGSSAHS  
ILLIANIPERFYODTISALNFTARSKEVINRPTNESLOPHALA PVKLSQKELLGS  
EAKAKAPESSTSPSPASQSLSLQKLSNMDPAMLENLUSMERLLGSSQS  
QGPLLTPRRERVMKTVBEKNLEIRLKKQKELEAKVLAQEPDPREKENTPT  
LOPPASVYGAFLKKAIVMPQRIQKRESSNOIQLLKKPKRKLPEPSEAVEK  
DEDYWEVQISPELLAHKRLDLLNKGSRRLSLQRIQKKAQLIVGWRHLGPPS  
EVEDLEQVEGISGQVESFLKANLLSLAASQHSFS"

## ORIGIN

Query Match 72.9%; Score 1067.8; DB 10; Length 2086;  
Best Local Similarity 84.1%; Pred. No. 2.2e-274;  
Matches 1204; Conservative 0; Mismatches 227; Indels 0; Gaps 0;  
QY 6 TCCTGTCGCTAGCAGATGAGCTACTCGTCGTCACCTCCAGCTCGCGTAAGCGT 65  
DB 82 TGGCCGCTGTGAAGCAAGAGAGCGCTTGGCCGCGTCCGCTCTAGCTCGAGTAAGGGT 141  
QY 66 GGCTGTGCGACTGCGGCGCATTTGGATGGAACAGCGGAGCAGTATCCCCCTGTGT 125  
DB 142 GGCTGTGCGGCTGCGCCGTTTATGGATGAGAGAGAGAGAGCTCCCTGTGT 201  
QY 126 CGCGGGCATGAGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAACCCAGGAGACTCT 185  
DB 202 CCGAGCCATAGACAGCTGCTCTTTGAAGTGGCTAACTGGAAGAAATACCCAGGAGCGCT 261  
QY 186 CAATAACAGTTTATGATCCCTCTATGGGAGAGAGAGTACTCAGCAGACATCTATGACGG 245  
DB 262 CAATAATACGTTTATGATCCCTCTATGGGAGAGAGAGACTCAGCAGGAACTATGTAGG 321  
QY 246 TTCACTGACGCCATCTTAAGGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGCCCTA 305  
DB 322 TTCACTGACGCCATCTTAAGGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGCCCTA 381  
QY 306 TGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCGCCAGCAGCAACCTGGGGT 365  
DB 382 TGGGCTTACTGGGCGACGGAGACACACAATGCTGGGAGCGCCAGCAGCAACCTGGAGT 441  
QY 366 GATCCCGCGGGCTCTCATGTAGACTCTCTGAGCTCACAAGGGAGAGGGTCCCGAGGCGG 425  
DB 442 GATTCTCTCGGGCTCTCATGTAGACTCTCTGCACTAGCAAGGGAGAGAGTCTGAGGGACG 501

QY 426 GCCATGGGCCCCCTTCTGTCAACATGTCTTACCTAGAGATCTACACAGGAGAAAGGTATTAGA 485  
DB 502 GCCATGGGACGTTTCTGTCTGTCTTCTTATTTAGAGATCTACACAGGAAAGGTATTAGA 561  
QY 486 CCTCTCGGACCTCTCTCCAGAGACCTGTGTAATCCGAGAAAGTGTCCGCGGGGAATATCT 545  
DB 562 CCTCTTGGATCTCTGATCAGGAGACCTGTGATCCGGAAGACTGCGCGAGGAAACATCT 621  
QY 546 GATTCCGGGTCTCTCCAGAGACCTCATCAGTAGCTTTGCTGATTTTGGAGCGGCACTTCT 605  
DB 622 GATCCAGGCTCTACACAGAGCCCATCACAGCTTCTCTGACTTCGAGAGCAGCTTCTCT 681  
QY 606 GCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGAGGCTCTCTCCCGCAG 665  
DB 682 TCCAGCCAGTAGAAATCGAGCTGTAGGAGCCACCGGCTTAAACAGCGCTCTTCCCGTAG 741  
QY 666 TCATGCTGTGCTCTCTGTCAGAGTGGACACGCGGGAAGCTTTGGTGGCCCATTTTCGCGACG 725  
DB 742 TCACGAGTGTCTCTTGGTCAAGGTAGATCAGCGTGAACGTTTGTACTCTCATTTCCGACG 801  
QY 726 AGAGGGAAGAACTCTACCTGATTGACTTGGCTGGGTTCAGAGACAAACCGCGCACAGGCAA 785  
DB 802 GGAAGGAAAACTCTACCTTATTGATTGGCTGGGTTCAGAGACAAACCGTCGACAGGCAA 861  
QY 786 CAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAA 845  
DB 862 CCAGGGCATTTGGCTCAAGAGAGCGGAGCCATCAACACCTCCCTCTTTGTACTTGGGCAA 921  
QY 846 AGTGTAGATGCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAC 905  
DB 922 AGTGTGTGATGATCAATAACAGGCGCTCCCTCGCATACCATACCGGAGACAGCAAGCTCAC 981  
QY 906 TCGCTATTTCGAGGACTCTCTGGGTGCTCAGCCCAAGTATCTTATTTGCAACATGTC 965  
DB 982 TCGTCTGCTCAGGACTCTCTGGGAGGCTCAGCTCATAGCATCTCTTATTTGCAACATGTC 1041  
QY 966 CCCTGAGAGACGCTTCTTACCTAGACAGTCTCCGCACTCAACTTGTCTGCGACGTCGCA 1025  
DB 1042 TCTGGAAGACGTTTTTACCAGATCAATTTTACGATTTGACATTTCTACTGCTAGGTCCA 1101  
QY 1026 GAGGTGATCAATCGGCTTTTATCAATGAGAGCTGCGAGCTCTCATGCTTGGAGCTGT 1085  
DB 1102 GAGGTGATTAACCGGCTTTTCAATGAGAGTCTACAGCTCATGCTTGGCAGCTGT 1161  
QY 1086 TAAGCTGTCTCAGAAAGAAATTTGCTTGTCCACAGAGGCAAGAGACCCGAGGCGCTTGA 1145  
DB 1162 TAAGCTGTCTCAGAAAGAAATTTGCTTGTAGGCCCTCAGAGGCAAGAGGCGCTTGA 1221  
QY 1146 GGAAGAGAGATTTGGGAGCGCTGAGCCATGCGCAGCTTCCAGCTCTGCTTCCAGAAACT 1205  
DB 1222 AGAAGATCAATGCGGCGCTGAGTCTACAGAGCTTCCAGCTCTGCTTCCAGAAACT 1281  
QY 1206 CAGCCCCCTTACAGAAAGCTAAGCAGCATGGACCCCGGCACTGCTGAGCGCTCTCTCAGCTT 1265  
DB 1282 AAGCTCTTACAGAAAGCTAAGCAATATGGATCCAGCCATGTTGGAGAACTCTCTGAGTAT 1341  
QY 1266 GAGCGCTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAGAGCG 1325  
DB 1342 GGAACGTTTCTGGGTTCCAGGGAAAGCCAAAGGAGCCCTTACTGTAATACCCCAAGAGCG 1401  
QY 1326 AGAGCGGATGGTCTTAAAGACAGTAGAAGAGAGAGGACCTAGAGATTTCAGAGGCTTAA 1385  
DB 1402 AGAACGATGGTCTCATGAAGACAGTGGAGGAAAGAACTTAGAGATTTCAGAGGCTTAA 1461  
QY 1386 GACCAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGAGCTCAGGA 1436  
DB 1462 GATGAAGCAAAAAGAACTGGAAGCCCAAGGTGTTGGCCCAAGAGGCGCACCGGA 1512

RESULT 10  
AR210054  
LOCUS

AR210054 1041 bp DNA linear PAT 20-JUN-2002

```

DEFINITION Sequence 5 from patent US 6387644.
ACCESSION AR210054
VERSION AR210054.1 GI:21512185
KEYWORDS
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1041)
AUTHORS Beraud, C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 5 14-MAY-2002;
FEATURES
    Location/Qualifiers
        1..1041
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
    Query Match 70.9%; Score 1038.4; DB 6; Length 1041;
    Best Local Similarity 99.9%; Pred. No. 1.6e-266;
    Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGCCACCTCCAGCTCGGTA 60
Db 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGCCACCTCCAGCTCGGTA 60
QY 61 AGGGTGGCTGTGGGACTCGGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
Db 61 AGGGTGGCTGTGGGACTCGGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
QY 121 TGTGTCGGGGCATGACAGCTCTCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
Db 121 TGTGTCGGGGCATGACAGCTCTCTCTAGAGATTGCTAACTGGAGGAACACCAAGAG 180
QY 181 ACTCTCAAAATPACAGTTTATGATGCTTCTATGGGAGAGGAGTACTCAGCAGACATCTAT 240
Db 181 ACTCTCAAAATPACAGTTTATGATGCTTCTATGGGAGAGGAGTACTCAGCAGACATCTAT 240
QY 241 GAGGTTTCAGTCAGCGCCATCTTAAGGCATCTGCTGGAAGGGCAGATGCCAGTGTCTT 300
Db 241 GAGGTTTCAGTCAGCGCCATCTTAAGGCATCTGCTGGAAGGGCAGATGCCAGTGTCTT 300
QY 301 GCCTATGGACCCACAGAGCTGGGAAGAGCAGACAACTGTCGGGAGCCGCCAGAACCT 360
Db 301 GCCTATGGACCCACAGAGCTGGGAAGAGCAGACAACTGTCGGGAGCCGCCAGAACCT 360
QY 361 GGGGTGATCCCGGGCTCTCATGACCTCTCTGACCTCACAAGGAGAGGGTGCAGAG 420
Db 361 GGGGTGATCCCGGGCTCTCATGACCTCTCTGACCTCACAAGGAGAGGGTGCAGAG 420
QY 421 GGGGGGATCCCGGGCTCTCATGACCTCTCTGACCTCACAAGGAGAGGGTGCAGAG 480
Db 421 GGGGGGATCCCGGGCTCTCATGACCTCTCTGACCTCACAAGGAGAGGGTGCAGAG 480
QY 481 TTAGACTCTCGACCTCTCTCGGAGACCTGCTAATCCGAGAAGACTGCCGGGGGAAT 540
Db 481 TTAGACTCTCGACCTCTCTCGGAGACCTGCTAATCCGAGAAGACTGCCGGGGGAAT 540
QY 541 ATCCTGATCCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
Db 541 ATCCTGATCCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTGAGCGGCAC 600
QY 601 TTCTGCGCAGCTCGAATTCGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCC 660
Db 601 TTCTGCGCAGCTCGAATTCGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCC 660
QY 661 CGCAGTATCTGCTCTCTGCTTCAAGGTGGACAGCGGGAAGCTTTGGCCCCCATTTCCG 720
Db 661 CGCAGTATCTGCTCTCTGCTTCAAGGTGGACAGCGGGAAGCTTTGGCCCCCATTTCCG 720
QY 721 CAGCGAGAGGGAAGAACTCTACTGATTTGACTTGGCTGGGTTCAGAGGACAAACCGGCGACA 780
Db 721 CAGCGAGAGGGAAGAACTCTACTGATTTGACTTGGCTGGGTTCAGAGGACAAACCGGCGACA 780

```

```

QY 781 GGCAAAAGGGCTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCCTGTTTCTCTG 840
Db 781 GGCAAAAGGGCTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCCTGTTTCTCTG 840
QY 841 GGCAAAAGTGGTAGATGCGCTGAATCAGGGCCTCCTCGTGTACCTTATCGGACAGCAAG 900
Db 841 GGCAAAAGTGGTAGATGCGCTGAATCAGGGCCTCCTCGTGTACCTTATCGGACAGCAAG 900
QY 901 CTCACCTGCGCTATTTCAGGACTCTCTGGTGGCTCAGGCCACAGTATCTTATTCGCAAC 960
Db 901 CTCACCTGCGCTATTTCAGGACTCTCTGGTGGCTCAGGCCACAGTATCTTATTCGCAAC 960
QY 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAGG 1020
Db 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAGG 1020
QY 1021 TCCAAGGAGGTGATCAATCG 1040
Db 1021 TCCAAGGAGGTGATCAATCG 1040

RESULT 11
AR210052
LOCUS Sequence 1 from patent US 6387644.
DEFINITION AR210052 1115 bp DNA linear PAT 20-JUN-2002
ACCESSION AR210052
VERSION AR210052.1 GI:21512182
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1115)
AUTHORS Beraud, C.
TITLE Motor proteins and methods for their use
JOURNAL Patent: US 6387644-A 1 14-MAY-2002;
FEATURES
    Location/Qualifiers
        1..1115
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
    Query Match 70.7%; Score 1035.4; DB 6; Length 1115;
    Best Local Similarity 99.9%; Pred. No. 1e-265;
    Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCTGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAGG 63
Db 78 GGTGCTGTGCTGCTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAGG 137
QY 64 GTGGCTGTGCGACTCGGGCCATTTCGGATGGAACAGCGGGAGCAAGTATCCCCCTCT 123
Db 138 GTGGCTGTGCGACTCGGGCCATTTCGGATGGAACAGCGGGAGCAAGTATCCCCCTCT 197
QY 124 GTGGCGGGCATCGACAGCTGCTCTCTAGAGATTGCTAATCGAGGAACACCAAGAGACT 183
Db 198 GTGGCGGGCATCGACAGCTGCTCTCTAGAGATTGCTAATCGAGGAACACCAAGAGACT 257
QY 184 CTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
Db 258 CTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 317
QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGATGCCAGTGTGCTGCC 303
Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGGAAGGGCAGATGCCAGTGTGCTGCC 377
QY 304 TATGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGCAGCAACCTGGG 363
Db 378 TATGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCGCAGCAACCTGGG 437
QY 364 GTGATCCCGGGCTCTCATGACCTCTCTGCAAGTCTCAGGAGAGAGGTCGCGAGGC 423
Db 438 GTGATCCCGGGCTCTCATGACCTCTCTGCAAGTCTCAGGAGAGAGGTCGCGAGGC 497

```

```
QY 424 CGGCCATGGGCCCTTCTGTGTCACGATGCTTACCTAGAGATCTACCGAGGAAGATATTA 483
Db 498 CGGCCATGGGCCCTTCTGTGTCACGATGCTTACCTAGAGATCTACCGAGGAAGATATTA 557
QY 484 GACCTCTGGACCTGCTTCGGAGACCTGCTATCCGAGAGACTGCGGGGGAATATC 543
Db 558 GACCTCTGGACCTGCTTCGGAGACCTGCTATCCGAGAGACTGCGGGGGAATATC 617
QY 544 CTGATTCGGGTCTCTCCAGAAGCCCATCAGTAGCTTGTGCTGATTTTGAGCGGCACATTC 603
Db 618 CTGATTCGGGTCTCTCCAGAAGCCCATCAGTAGCTTGTGCTGATTTTGAGCGGCACATTC 677
QY 604 CTGCAGGACGTCGAATCGGACTGTAGAGACCCCGGCTCAACACGAGCTCTCTCCGCG 663
Db 678 CTGCAGGACGTCGAATCGGACTGTAGAGACCCCGGCTCAACACGAGCTCTCTCCGCG 737
QY 664 AGTCATGCTGTCTCTGCTCAAGTGGACGAGCGGGAACGTTTGGCCCAATTTTCGCAG 723
Db 738 AGTCATGCTGTCTCTGCTCAAGTGGACGAGCGGGAACGTTTGGCCCAATTTTCGCAG 797
QY 724 CGAGAGGGAACCTCTACCTGATTTGACTTGGCTGGGTGAGAGACCAACCGGCGCACAGGC 783
Db 798 CGAGAGGGAACCTCTACCTGATTTGACTTGGCTGGGTGAGAGACCAACCGGCGCACAGGC 857
QY 784 AACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGCG 843
Db 858 AACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGCG 917
QY 844 AAGTGTGTAGTGCCTGATCAGGCGCTCCCTGCTGCTACCTTATCGGACAGCAAGCTC 903
Db 918 AAGTGTGTAGTGCCTGATCAGGCGCTCCCTGCTGCTACCTTATCGGACAGCAAGCTC 977
QY 904 ACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTCTTATTCGCAACATT 963
Db 978 ACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTCTTATTCGCAACATT 1037
QY 964 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTGTGTCGAGTCC 1023
Db 1038 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGACCTCAACTTGTGTCGAGTCC 1097
QY 1024 AAGGAGTGTATCATCG 1040
Db 1098 AAGGAGTGTATCAATTG 1114

RESULT 12
AC101919/c
LOCUS AC101919 151041 bp DNA linear HTG 14-MAR-2003
DEFINITION Mus musculus clone RP24-158A6, WORKING DRAFT SEQUENCE, 4 unordered
pieces.
ACCESSION AC101919
VERSION AC101919.4 GI:28951338
KEYWORDS HTG; HTGS, PHASE1; HTGS, DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 151041)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus, clone RP24-158A6
Unpublished
2 (bases 1 to 151041)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, I., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
```

McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Nobus, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 151041)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,  
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepe, I.,  
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,  
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,  
Graham, L., Grand-Pierre, N., Hagaz, N., Hagopian, D., Hagos, B.,  
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C.,  
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,  
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (14-MAR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 14, 2003 this sequence version replaced gi:28631351.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L17745  
Center clone name: 158 A.6  
----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 150558 bases at least Q40  
Consensus quality: 150654 bases at least Q30  
Consensus quality: 150687 bases at least Q20  
Insert size: 151000; agarose-fp  
Insert size: 150741; sum-of-contigs  
Quality coverage: 11.8 in Q20 bases; agarose-fp  
Quality coverage: 11.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 4 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.  
1 33517: contig of 33517 bp in length  
33518 33617: gap of 100 bp  
33618 42736: contig of 9119 bp in length



Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Coyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., LaRocque, K., Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Melldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Robert, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 164759)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cooke, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, V., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 164759)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cooke, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 23, 2004 this sequence version replaced gi:45504330.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
----- Project Information  
Contact: sequence\_submissions@broad.mit.edu  
Center project name: L17344  
Center clone name: 344\_C\_18  
-----

FEATURES	source	Location/Qualifiers
repeat_region	1..164759	/organism="Mus musculus"
misc_feature	1..1655	/mol_type="genomic DNA"
		/db_xref="taxon:10090"
		/chromosome="1"
		/map="1"
		/clones="RP24-344C18"
		/clone_lib="RPCI-24 Male Mouse BAC"
		/rpt_family="L1_MM"
		/note="clone boundary"
		clone_end:SP6
		site:MboI
repeat_region	1767..1787	/rpt_family="AT_rich"
repeat_region	2115..2157	/rpt_family="AT_rich"
repeat_region	2236..2258	/rpt_family="AT_rich"
repeat_region	3252..3420	/rpt_family="Lx7"
repeat_region	3464..3588	/rpt_family="Lx7"
repeat_region	3605..3666	/rpt_family="CT-rich"
repeat_region	3673..3731	/rpt_family="(TTC)n"
repeat_region	3732..3911	/rpt_family="(TTC)n"
repeat_region	4000..4180	/rpt_family="CT-rich"
repeat_region	4202..4363	/rpt_family="Lx7"
repeat_region	4529..4765	/rpt_family="MTB"
repeat_region	7234..7442	/rpt_family="B3"
repeat_region	8445..8509	/rpt_family="T-rich"
repeat_region	complement(8512..8591)	/rpt_family="MLT11"
repeat_region	9336..9632	/rpt_family="MLT1J"
repeat_region	complement(9800..11396)	/rpt_family="L1_MM"
repeat_region	complement(11397..17527)	/rpt_family="L1P"
repeat_region	complement(18235..18777)	/rpt_family="Lx"
repeat_region	18775..18890	/rpt_family="Lx"
repeat_region	19260..19949	/rpt_family="RMER19A"
repeat_region	20118..20297	/rpt_family="MLT1D"
repeat_region	complement(20386..20639)	/rpt_family="Lx2"
repeat_region	20634..20839	/rpt_family="Lx4"
repeat_region	20876..21793	/rpt_family="RLTR13D"



```

repeat_region 21799..21907
/rpt_family="RMER13A"
repeat_region 21918..22433
/rpt_family="Lx4"
repeat_region 22591..22651
/rpt_family="(TG)n"
repeat_region 22733..24486
/rpt_family="L1_MM"
repeat_region 24494..24850
/rpt_family="L1_MM"
repeat_region 27139..27166
/rpt_family="AT_rich"
repeat_region 27737..27834
/rpt_family="MLT1H"
repeat_region 28328..28506
/rpt_family="MLT1K"
repeat_region 28513..28541
/rpt_family="AT_rich"
repeat_region 28813..28858
/rpt_family="AT_rich"
repeat_region 29627..30186
/rpt_family="L1_MM"
repeat_region 30465..30578
/rpt_family="T-rich"
repeat_region complement(30581..34457)
/rpt_family="L1_MM"
repeat_region complement(34455..35895)
/rpt_family="L1"
repeat_region complement(36134..36736)
/rpt_family="Lx4"
repeat_region 36749..36920
/rpt_family="Lx2"
repeat_region 37617..37751
/rpt_family="(TA)n"
repeat_region 37958..37986
/rpt_family="(GAA)n"

```

Query Match 60.2%; Score 881.8; DB 10; Length 164759;

Best Local Similarity 78.4%; Pred. No. 1.8e-224;  
Matches 1116; Conservative 0; Mismatches 282; Indels 25; Gaps 4;

```

QY 12 TCGGCTAAGCAAGATTGGAGCTACTCGTCTGCACCTCCAGCTCGGCTAGGGTGGCTGT 71
DB 146111 TTGATTAAGCAAGTAGGGATTGGCTGCTCTGCTTCTAGCTTGGATAGGGTGGCTAT 146052
QY 72 GCGACTCGGCGCAATTGTGGATGGAAACAGCGGAGCAAGTGAATCCGCCCTGTGTGGGG 131
DB 146051 ATAATTGCTGCTTTGTGGATGAAGACAGAGCAAGGAGCAACCCCTGTGTACGAGC 145992
QY 132 CATGGAAGCTGCTCTCTAGAGATTGCTTAAGTGGAGAACCAACAGGAGACTCTCAANTA 191
DB 145991 CATAGACAGCTGCTCTCTTGAAGTGGCTAACTGGAGGAATAACAGGAGACTATCAANTA 145932
QY 192 CCAGTTTATGCTCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCAAGTTTCAGT 251
DB 145931 TCAGTTTATGCTCTTCTATGGGAGAGAGACTCAGTGGACATCTATGAGTTTATG 145872
QY 252 GCAGCCCATCTTAAGGCACTTGTGTGAAGGCGAGAATGCCAGTGTGTTGCTTATGAGCC 311
DB 145871 ACAGCCCATCTTAAGGCACTTGTGTGAAGGCGAGAATGCCAGTGTGTTGCTTATGAGCC 145812
QY 312 CACAGGAGCTGGAGAGCGCACAACTGCTGGGAGCCAGCAACCTGGGGTGTATCCC 371
DB 145811 GACTGGGCGAGGGAAGA--CACACAAATGTTGGGAGCCAGAACAGCTGAGATGATCC 145754
QY 372 GCGGCTCTCATGAGACCTCTCTGAGCTCAAGGAGAGAGGGTGCAGAGGCGCGCCCATG 431
DB 145753 TCAGACTCTCCAGACCTCTCTGACCTTAACAGGGAAGAGAACTCTGAGGCTGCGCCATG 145694
QY 432 GGCCCTTTCTGTACCACTGCTTACCTTAGAGATCTACAGAGAGAGGATATTAGACCTCCT 491
DB 145693 GGACATTTCTGTGCTTGTGCTATTATTAGATCTCTTAACAGGAATAAGTATTAGACCTCTT 145634

```

```

QY 492 GGACCTGCTTCGGGAGACCTGGTAATCCGAGAGAGACTGCGGGGGAATATCTCTGATTC 551
DB 145633 GGATCCTGCAATCAGAAGACCTAGTGTTCGAGAAAGGCTGACGAGGAAATCTCTGATTC 145574
QY 552 GGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGTGATTTTGTAGCGGCACTTCTTCGCGAGC 611
DB 145573 AGGCTCACAGAAAGCCCATCAGTAGTTTCTCTGACTTCGAGAAAGCACTTCTCTTAGT 145514
QY 612 CAGTCGAAATCGGACTGTAGAGCCACCGGCTCAACAGGCTCCTCCGAGCTCATGC 671
DB 145513 CAGTAGAAATCGAGTTGTAGGAGCTACTCTGCTTTAACGAGGCTCTTGCGGCTAGTCAATGC 145454
QY 672 TGTGCTCTGCTCAAGTGGACAGCGGGAACGTTTGGCCCCATTTGGCCAGCAGAGAGGG 731
DB 145453 AGCGCTCTTGGTCAAGGTAGATCAGCATGAACGTTTGAATCCATTTCCGAGTGGGAGG 145394
QY 732 AAAACTCTACCTGATTTGACTTTGGCTGGGTAGAGGACAAC--GGGCGACAGGCAACA 787
DB 145393 AAAACTCTACCTTATTGATTTGGCTGGTTAGAGGACAACCGTCAAGTCCGACAGGCAACC 145334
QY 788 AGGCGCTTCGGCTAAAGAGAGTGGGCCATCAACACCTCCCTGTTTCTCTGGGCAAG 847
DB 145333 AGGGCAITCTGCTCAAGAGAGTGGAGCATCAACACCTCCCTCTTTGTACTGGGCAAG 145274
QY 848 TGTAGATGCGCTGAATCAGGGCTCTCCCTCGTGTACCTTATCGGACAGCAAGCTCATCTC 907
DB 145273 TAGTGATGCAATTAACCAAGGCTCTCTCATATACCATACAGGACAGGAGCTCATCTC 145214
QY 908 GCCTATTCCAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCCACATTTGCC 967
DB 145213 GTCTGTGCAAGACTCTCTGAGAGGCTCAGCTCATAGCATCTCTATTGTACGATTTGCTC 145154
QY 968 CTGAGAGAGCTTCTTACCTAGACACAGTCTCGGCACTCAACTTTGCTGCGAGGCTCAAAG 1027
DB 145153 CTGAAGATGTTTTCCTGCTGATACATCTAGCATTAACCTTCTACTGCTAGTCCAGG 145094
QY 1028 AGGTGATCAATCGGCTTTTACCAATGAGAGCTCAGCCCTCATGCTTGGAGCTCTGTTA 1087
DB 145093 AGGTGATTAACCAAGGCTTTTCAACCAAGGAGTGTACAACCTCAGGCTTGACACCTGTTA 145034
QY 1088 AGCTGTCTCAGAAAGTGTCTTGTCCACAGAGGCAAGAGAGCCGAGGCGCTCAGG 1147
DB 145033 AGCTGTCTCAGAAAGTGTCTTGTCTTATCAGAGGCAAGAAAGCCAAAGGCTCAG 144974
QY 1148 AAGAGAGATTCGGAGCCCTGAGCCATGCGGAGCTCCAGGCTCTGCTCTCCAGAACTCA 1207
DB 144973 AAGATCAACTGGAGCCCTGAGTCTTAGCATCTCCAGCATCTGCTTCCAGAAA--- 144918
QY 1208 GCGCCCTACAGAAAGCTAAGCAGCATGGACCCCGGCAATGTGGAGCGCTCTCAGCTTGG 1267
DB 144917 -----CTAAGCAATATGGATCTAGCCATGTGGGAAAGCTCTCTGAGTATGG 144872
QY 1268 ACCTGCTGCTTCCAGGAGGAGCCAGGCGCCCTCTGTTGAGTACCCCAAGGAG 1327
DB 144871 AATGTTTGTGGTTCACAGGAGGAGCCAGAGAACCTCTACTGATATACCCCAA--CGAG 144813
QY 1328 AGCGGATGTGTCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTTGAGAGCTTAAGA 1387
DB 144812 AGCGAATGTGTCTAATGAAGCGTGGAGGACAGATATTTGAGATTTGAGAGCTTAAGA 144753
QY 1388 CGAAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCCGCAAGGC 1430
DB 144752 TGAAGCAAAAGAACTGGAGGCGCAGGTGCTGGTCCAGGAAGC 144710

```

# RESULT 14

AC101908/c

LOCUS

DEFINITION

PROGRESS \*\*\*

AC101908

AC101908.4

VERSION

KEYWORDS

AC101908 297639 bp DNA linear HTG 18-JUN-2004  
Mus musculus chromosome 1 clone RP24-275J1 map 1, \*\*\* SEQUENCING IN  
PROGRESS \*\*\* 8 unordered pieces.  
AC101908  
AC101908.4 GI:44199137  
HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.



SOURCE  
ORGANISM

Mus musculus (house mouse)

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 297639)

Birren, B., Nusbaum, C. and Lander, E.

Mus musculus chromosome 1, clone RP24-275J1

JOURNAL

REFERENCE  
AUTHORS

Unpublished

2 (bases 1 to 297639)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Canarata, J., Campiano, A., Chang, J., Chazaro, B.,  
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooper, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,  
Lamaze, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McSwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigliio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 297639)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V.,  
Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J.,  
Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B.,  
Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,  
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,  
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hegopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,  
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,  
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,  
Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,  
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,  
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,  
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,  
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,  
Zimmer, A. and Zody, M.

Direct Submission

Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Feb 27, 2004 this sequence version replaced gi:31880232.

All repeats were identified using RepeatMasker:

Smith, A. F. A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@broad.mit.edu](mailto:sequence_submissions@broad.mit.edu)

----- Project Information

Center project name: L17732

Center Clone name: 275\_J\_1

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 63747: contig of 63747 bp in length  
63848 63847: gap of 100 bp  
63848 73892: contig of 10045 bp in length  
73893 73992: gap of 100 bp  
73993 154450: contig of 80458 bp in length  
154451 154550: gap of 100 bp  
154551 167808: contig of 13258 bp in length  
167809 167908: gap of 100 bp  
167909 167911: contig of 8803 bp in length  
167912 176811: gap of 100 bp  
176812 233500: contig of 56689 bp in length  
233501 233600: gap of 100 bp  
233601 241222: contig of 7622 bp in length  
241223 241322: gap of 100 bp  
241323 297639: contig of 56317 bp in length.

FEATURES  
source

1. 297639  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clones="RP24-275J1"  
/clone\_lib="RPCI-24 Male Mouse BAC"

ORIGIN

Query Match	50.6%;	Score	740.6;	DB 2;	Length	297639;
Best Local Similarity	78.2%;	Pred. No.	1.1e-186;			
Matches	916;	Conservative	0;	Mismatches	249;	Gaps 2;

  

Qy	12	TCGGCTAAGCAAGATTGGAGCTACTCTGTCGTCACCTCCAGCTCGCGTAAGGGTGGCGT	71
Db	1179	TTGTATTAAGCAAGTAGGATTGGCTGCTGCTCTTAGCTGAGTAAGGGTGGCTAT	1120
Qy	72	CGGACTCGGCCCAATTTCTGGATGAAACAGCGGGAGCAAGTGAATCCCCCTGTGTGCGGGG	131
Db	1119	ATAATTGTGCTGTTTGTGGATGAAGACAGAAAGCAAGAGCACCCTGTGTACGAGC	1060
Qy	132	CATGGACAGCTGCTCTCTAGAGATTGTAAGTGGAGAACCCAGGAGACTCTCAATA	191
Db	1059	CATAGACAGCTGCTCTCTTGAAGTGGCTAACTGGAGGAAATACCAAGAGACTATCAATA	1000
Qy	192	CCAGTTTGTATGTCCTTCTATGGGAGAGGAGTACTCAGCAGACATCTATGCAAGTTCAGT	251
Db	999	TCAGTTTGTATGTCCTTCTATGGGAGAGAGACATCTAGTGGACATCTATGAGTTTGT	940
Qy	252	GCAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTTATGGAC	311
Db	939	ACAGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTTATGGAC	880
Qy	312	CACAGGAGCTGGGAGAGCAGACATCTGTTGGGAGCCCGCAGAGCACTGGGGTATCC	371
Db	879	GACTGGGGCAGGGAAGA--CACAAATGTTGGGAGCCCGCAGAGCACTGGAGTATTC	822
Qy	372	CGCGGCTCTCATGGACCTCTGAGCTCAACAAGGAGGAGGGTCCCGAGGGCCGCCATG	431
Db	821	TCAGACTCTCCAGACCTCTGACCTTAAAGGGAAGAGAACTCTGAGGGCTGGCCATG	762
Qy	432	GGCCCTTCTGTGTGCTTACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTCT	491
Db	761	GGACATTTCTGTTGCTTTTGTGCTATTATTAGATCTCTACAGGAAATAGTATTAGACCTCT	702
Qy	492	GGACCTGCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGGAATATCTCTGATTC	551
Db	701	GGATCTCTTATATAGTGTACCTAAATGTCAGCGCCAGGAGGAGAAATCTCTGATTC	642

552 GGGTCTCTCCAGAGCCCATCAGTAGCTTGTCTGATTTTGAGCGGCACTCTCCTGCCAGC 611  
 641 AGGCTCTCACAGAGCCCATCAGTAGTTTCTGACTTCGAGAGCACTCTCTTCTAGT 582  
 612 CAGTGGAAATCGAGTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGCAGTCATGC 671  
 581 CAGTGGAAATCGAGTGTAGGAGCTACTCTGCTTAACACAGCGCTCTTGCCTGAGTCATGC 522  
 672 TGTGCTCTGCTGCTCAAGTGGAGCGGGAAGCTTTGGCCCCCATTTGCGCAGCAGAGGG 731  
 521 AGGCTCTTGGTCAAGTAGATCAGATGAAGTTTGACTCAATTCGCCAGTGGGAAGG 462  
 732 AAAACTCTACCTGATGACTTGGCTGGGTGAGAGCAAC-----GGCGCACAGGCAACA 787  
 461 AAAACTCTACCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 402  
 788 AGGCGCTTGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGCGCAAG 847  
 401 AGGCGCTTGGCTCAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGCGCAAG 342  
 848 TGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907  
 341 TAGTGTGATGATTAACCCAGGCTCTCTCATATACCATACAGGAGCAGAGGCTCACTC 282  
 908 GCTATTGAGGACTCTCTGGGTGGTCAAGCCACAGTATCTTATTCGAGCAAGCACTCACTC 967  
 281 GTCTGTTGCAAGACTCTCTGAGAGGCTCAGCTCATAGCATCTCTCATTTGTCAGCATTTGCTC 222  
 968 CTGAGAGAGCTTCTACCTAGACACAGTCTCGGCACTCACTTGTCTGCGAGGCTCAAGG 1027  
 221 CTGAAGAGTGTCTTACCTGGATACATCTCAGCATTAACCTCACTGCTAGTGTCAAGG 162  
 1028 AGGTGATCAATCGGCTCTTTTACCAATGAGAGCTCGAGCTCATGCTTTGGGACCTGTGTA 1087  
 161 AGGTGATCAATCGGCTCTTTTACCAAGAGGAGTGTCAACCTCAGGCTTGTGACACCTGTGTA 102  
 1088 AGCTGTCTGAGAAGATTTCTTGGTCCAGCAGCAAGAGAGCCGAGGCGCTTGTGAGG 1147  
 101 AGCTGTCTGAGAAGATTTCTTGGTCCAGCAGCAAGAGAGCCGAGGCGCTTGTGAGG 42  
 1148 AAGAGGAGATTTGGAGGCTTGGAGCCATGCG 1178  
 41 AAGATCACTGGAGGCTTGTGATCTATAGC 11

RESULT 15  
 AC102127 196674 bp DNA linear ROD 01-MAY-2004  
 LOCUS Mus musculus chromosome 1, clone RP23-202A19, complete sequence.  
 DEFINITION AC102127  
 ACCESSION AC102127  
 VERSION AC102127.11 GI:46931440  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 196674)  
 Mus musculus chromosome 1, clone RP23-202A19  
 Unpublished  
 2 (bases 1 to 196674)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campoliano,A., Chang,J., Chararo,B.,  
 Choepe,Y., Collamore,M., Collins,S., Collymore,A., Cooke,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Govette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 MacLean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rize,C., Rogov,P.,  
 Roman,S., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 196674)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
 Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,  
 Choepe,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,  
 Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
 Erickson,J., Fero,S., Ferreira,P., FitzGerald,M., Gage,D.,  
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
 Retta,R., Rize,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
 Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (10-APR-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 196674)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
 Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,  
 Choepe,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,  
 Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
 Erickson,J., Fero,S., Ferreira,P., FitzGerald,M., Gage,D.,  
 Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
 Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
 MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
 Retta,R., Rize,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
 Zimmer,A. and Zody,M.

Direct Submission  
 Submitted (01-MAY-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On May 1, 2004 this sequence version replaced gi:46358225.  
 All repeats were identified using RepeatMasker:  
 Smith,A.F.A. & Green,P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/MIT Center for Genome Research  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence submissions@broad.mit.edu  
 ----- Project Information

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

TITLE  
 JOURNAL  
 COMMENT

Center project name: L18071  
Center clone name: 202\_A\_19

FEATURES	Source	Location/Qualifiers
	1..196674	/organism="Mus musculus"
		/mol_type="genomic DNA"
		/db_xref="taxon:10090"
		/chromosome="1"
		/map="1"
		/clone="RP23-202A19"
		/clone_lib="RPCI-23 Female Mouse BAC"
misc_feature	1..4229	/note="wgs end extension"
repeat_region	1..193	/clone_end:SP6"
unsure	614..746	/rpt_family="MTA"
unsure	701..706	/note="single clone coverage"
repeat_region	1208..1291	/note="<30 qual single clone coverage"
repeat_region	1571..1842	/rpt_family="(CTA)n"
repeat_region		/rpt_family="RLTRETN MM"
repeat_region		complement(1974..2067)
repeat_region	3219..3319	/rpt_family="B4"
repeat_region		/rpt_family="CA-rich"
repeat_region	3388..3448	/rpt_family="(CA)n"
repeat_region	3626..3688	/rpt_family="(CAAA)n"
repeat_region	3925..4374	/rpt_family="Lx"
misc_feature	4230..4235	/note="clone boundary"
		clone_end:SP6
		site:ECORI"
repeat_region	4383..4411	/rpt_family="AT-rich"
repeat_region	4445..4496	/rpt_family="(TCTCTG)n"
repeat_region	4509..4638	/rpt_family="(TG)n"
repeat_region	4741..4784	/rpt_family="(TG)n"
repeat_region	4785..4832	/rpt_family="(TA)n"
repeat_region	4833..4955	/rpt_family="(TG)n"
repeat_region	5398..5431	/rpt_family="(GAAA)n"
repeat_region		complement(5434..5561)
repeat_region		/rpt_family="RMER15"
repeat_region		complement(6185..6272)
repeat_region		/rpt_family="RSINE1"
repeat_region	6516..6570	/rpt_family="(T)n"
repeat_region		complement(6571..6823)
repeat_region	6821..6939	/rpt_family="L1_MM"
repeat_region	7007..7285	/rpt_family="Lx5"
repeat_region	7292..7848	/rpt_family="Lx5"
repeat_region	8187..8286	/rpt_family="RSINE1"
repeat_region		complement(8823..9061)
repeat_region		/rpt_family="B4"
repeat_region		complement(9243..9815)
repeat_region		/rpt_family="L1"

repeat_region	9832..9861	/rpt_family="(A)n"
repeat_region	9862..10368	/rpt_family="Lx2"
repeat_region		complement(10373..11559)
repeat_region	11555..12196	/rpt_family="L1_MM"
repeat_region		/rpt_family="Lx2"
repeat_region		complement(13221..13381)
repeat_region	13437..13578	/rpt_family="B3A"
repeat_region	15222..15254	/rpt_family="B1_MM"
repeat_region		/rpt_family="AT-rich"
repeat_region		complement(15315..15605)
repeat_region		/rpt_family="RMER20"
repeat_region		complement(16008..16517)
repeat_region		/rpt_family="RMER5"
repeat_region	16673..16713	/rpt_family="(TG)n"
repeat_region	16809..16866	/rpt_family="(TG)n"
repeat_region		complement(16922..17052)
repeat_region	17101..17490	/rpt_family="B4"
repeat_region		/rpt_family="MLT2B1"
repeat_region	17725..17762	/rpt_family="(GAAA)n"

Query Match 48.3%; Score 706.8; DB 10; Length 196674;  
Best Local Similarity 73.3%; Pred. No. 1.2e-177;  
Matches 1013; Conservative 0; Mismatches 357; Indels 12; Gaps 8;

Qy	69	TGTGGCACTGCGGCCATTGTGTGATGGAACAGCGGGAGCAAGTATCCCCCTGTGTGGC	128
Db	101632	TTTATGACTATGCCCATTTATGATGAAGAAAAAGGAGGATGCCCTTGTGTAA	101691
Qy	129	GGCATTGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACAGAGACTCTCAA	188
Db	101692	AGGCATATAAAGCTGCTCTTGGAGTGGCCACTTGGAGAAATCTTATGAGATATA	101751
Qy	189	ATACCAATTTGATGCTTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCAGTTC	248
Db	101752	ATATCAGTTTAAATGTCTTTTATTGAGAGAAGACACTCAGCAGAACATCTAAGTAGTTC	101811
Qy	249	AGTGCAGCCCATCTTAAGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGCCTATGG	308
Db	101812	AGTACAGCCTATCTCTGGCACTTGTGGATGGGCAAGATACCAGTGTGCTTACATATGG	101871
Qy	309	ACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGGCC---AGAGCAACTGGGG	364
Db	101872	GACTATTGAGGAGGAGACACATATATCTGGGTAGCCCATGGGAACAACATGGAG	101931
Qy	365	TGATCCCGGGGCTCTCATGAGCTCTCTGAGCTCAAGGGAGGGGTGCCGAGGGCC	424
Db	101932	TGATTCCTTAGCCTCTCATAGACCTCTGCAGATAACAAGGGGAAGAGTCTGAAGGCT	101991
Qy	425	GGCCATGGGCCCTTTCTGTCCACCATGTCTTACCTAGAGATCTACCAGGAAGTATAG	484
Db	101992	GGCCATGGGACATTTCTGTGAGGATGTCTTATTAGAGGTTCTACCAGGAAGGATATAG	102051
Qy	485	ACCTCTCGGACCTGCTTGGGAGACTGTGTAATCCGAGAAGACTGCCGGGGGAATATCC	544
Db	102052	ACCTCTTATATCTGTGATCAGGAGACCTAGTATATGAGAGATTTGTCAGGGAGACATCC	102111
Qy	545	TGATTCGGGTCTCTCCAGAACCCCATCAGTAGCTTTTGTGATTTTGTAGCGGCACTTCC	604
Db	102112	TGATTCAGTTCTACCCAGGACCCCAATACTTGTCTCTGACTTTTGTAGCAGTACTT-C	102170
Qy	605	TGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGCA	664
Db	102171	TGCCAGCCTGTAGAAATCAAGTTGTAGGAGCTACTCTGCTTAAAGCAGTGTCTTACCATA	102230

QY 665 GTCATGCTGTCTCTCTGTGTCAGGTGGACCAAGCGGGAACGTTTGGCCCCCATTTTGGCCAGC 724  
 Db 102231 -TCATTCATGCTCTTTATCATAGGTGGATCAGCATGGACATTTGACTCTGTTCACCTAGT 102289  
 QY 725 GAGAGGAAAACCTCTACCTGATTTGCTGGCTGAGGACCAACCGGGCCAGGCA 784  
 Db 102290 GAAAAGAAAACCTCTACCATATTTGATTTGGTTGGCTCAGAGGCAACCAACA-AGGA 102348  
 QY 785 ACAAGGGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGCA 844  
 Db 102349 ACAGTGCAATTAACCTCAAGAGAGTGGGCCATCAACACCTGCTTC-TTGTACTGGCA 102407  
 QY 845 AAGTGTAGATGGCTGAATCAGGGCTCCCTGCTGATCCTTATCGGACAGCAAGCTCA 904  
 Db 102408 AAGTGGAGGTTCATCATATTAGGGCTCCCTCTATATTAATCTGGGACAGAAAATCA 102467  
 QY 905 CTCGCCATATGCAAGGCTCTCTGGGTGGCTCAGCCACAGTATCCTTATTTGCCACATTTG 964  
 Db 102468 TTTGCCATATGGCAGACTCTCTGGAGGGCTCAGCTCAATCAATCATCATTTACCAACTG 102527  
 QY 965 CCCTCAGAGAGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTGTCTCCAGGTCCA 1024  
 Db 102528 CTCCTGAAGACATTTTACTTTGATCATTTCTGACATTTGACCTTCACTGTAGGTCCA 102587  
 QY 1025 AGAGGTGATCAATCGGCCCTTTTCAATGAGAGCTCGAGCTCATGCTTGGGACCTG 1084  
 Db 102588 AGGAGCTGATTAATTTGGCCCTTTTCAACCAATGGAGTCTACAGCTCATGCTTACACACTG 102647  
 QY 1085 TTAAGCTGCTCAGAAAGATTTGCTTGGTCCACGAGCAAGAGAGCCCGAGGCCCTG 1144  
 Db 102648 TTAAGCTTTGGCAAAAGAACTTCTAGGCCATCAAGGCCAAGAAAGCAAGGTCTTG 102707  
 QY 1145 AGGAAGAGAGATTTGGGAGCCCTGAGCCCATGGCCTCCAGCTCTGCTCTCCAGAAAC 1204  
 Db 102708 TAGAAGNATCAATGGGGAGTCTTAACCTACAGCAGTTTTATCACTGCTCCAGAAAC 102767  
 QY 1205 TCAGCCCCCTACAGAGCTTAAGCAGCATGGACCCGGCCATGCTGGAGCGCTCTCAGCT 1264  
 Db 102768 TAAGCCTCTTATAGAACCTAAGCAATGTGGATCCAGTATGTTGGAGAGTTCCTGAGTT 102827  
 QY 1265 TGGACCGCTGCTTGGCTCCAGGGAGCCAG--GGGGCCCTCTCTGTGAGTACCCCAAA 1322  
 Db 102828 TATATCAATTTGTTGGGATCTCAGGGGACAAAGAGGTATATTCTACTGAATACCCCAAA 102887  
 QY 1323 GCGAGA-GCGGATGGTCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGGC 1381  
 Db 102888 GCAAGACAAGATGGTACTCATTAAGACAGT-GTAGGAAGAACTTTGGAGATTGAGAGAC 102946  
 QY 1382 TTAAGACGAACCAAAAGAACTGGAGGCCCAAGATTTGGCCCAAGGCTGAGGAAAAGG 1441  
 Db 102947 TTAAGGTGAAGTAGAAGAACTATAGGCAAGGTGTTGGCACAGGATCCTAGATATGA 103006  
 QY 1442 AG 1443  
 Db 103007 AG 103008

Search completed: November 10, 2004, 13:18:11  
 Job time : 6475.64 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 09:45:10 ; Search time 756.977 Seconds  
(without alignments)  
10152.424 Million cell updates/sec

Title: US-10-797-893-7  
Perfect score: 1464  
Sequence: 1 atgggtcgctgtcggtgaag.....accattgtccacaatgtga 1464

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0  
Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1464	100.0	1464	10	ADC23343
2	1464	100.0	1464	12	ADQ60233
3	1461	99.8	1538	10	ADC23339
4	1461	99.8	1538	12	ADQ60229
5	1458	99.6	2097	12	ADQ09241
6	1447	98.8	2099	10	ADK40995
7	1443.2	98.6	2104	3	AAL15853
8	1038.4	70.9	1041	10	ADC23341
9	1038.4	70.9	1041	12	ADQ60231
10	1035.4	70.7	1115	10	ADC23337
11	1035.4	70.7	1115	12	ADQ60227
12	480	32.8	491	6	ABK70260
13	376.4	25.7	386	4	AAS38879
14	265	18.1	464	9	ACH44493
15	233.2	15.9	531	12	ACH77408
16	230.4	15.7	232	12	ACH91108
17	185.6	12.7	2095	11	ADM01564
18	185.6	12.7	3624	12	ADQ44181
19	156.8	10.7	3374	10	ADJ95073
20	154.8	10.6	2375	8	ABX34596
21	138.6	9.5	4108	6	ABA94614

22	138.6	9.5	4108	8	ABS57215
23	137.2	9.4	2675	10	ADJ95077
24	136.6	9.3	1152	6	ABA94616
25	136.6	9.3	1152	8	ABS57216
26	128.8	8.8	897	4	AAI70088
27	128.8	8.8	897	6	AAI39619
28	128.8	8.8	897	12	ADM81051
29	127.8	8.7	488	5	ABV47991
30	127.6	8.7	482	10	ABT40796
31	126.6	8.6	1026	6	ABQ73061
32	126.6	8.6	1026	9	AAL56806
33	126.6	8.6	1026	10	ADB66785
34	124.2	8.5	1011	6	ABQ73062
35	124.2	8.5	1011	9	AAL56807
36	124.2	8.5	1011	10	ADB66787
37	121.8	8.3	3661	6	AAD28568
38	121.8	8.3	3694	10	ADC10189
39	121.4	8.3	1014	6	ABA94615
40	121.2	8.3	3570	12	ADQ19012
41	119.8	8.2	2994	4	ABL28591
42	119.4	8.2	2034	4	ABL16733
43	119.4	8.2	4034	4	ABL16732
44	119.2	8.1	2132	10	ADC30338
45	117	8.0	1839	10	ADC30569

ALIGNMENTS

RESULT 1  
ADC23343  
ID ADC23343 standard; DNA; 1464 BP.  
XX  
AC ADC23343;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).  
XX  
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1464  
FT /tag= a  
FT /product= "KID protein"  
XX  
DN US6387644-B1.  
XX  
PD 14-MAY-2002.  
XX  
PF 28-NOV-2000; 2000US-00724224.  
XX  
PR 20-APR-1999; 99US-00295612.  
XX  
PR 20-JUN-2000; 2000US-00597292.  
XX  
XX (CVTO-) CYTOKINETICS INC.  
XX  
XX Beraud C;  
XX  
XX WPI; 2003-706919/67.  
XX  
XX P-PSDB; ADC23344.  
XX  
XX Identifying a candidate agent as modulator of function of a target  
XX protein for treating cellular proliferation disorders by adding a  
XX candidate agent to a mixture of the target protein that  
XX directly/indirectly produces ADP or phosphate.  
XX  
XX Disclosure; SEQ ID NO 7; 26pp; English.

XX This invention relates to a novel method for high throughput screening  
CC systems used to identify compounds for the treatment of cellular  
CC proliferation disorders. Specifically, it refers to candidate agents that  
CC are capable of modulating the activity of target proteins having motor  
CC domains, such that the target protein directly or indirectly produces ADP  
CC or phosphate. Furthermore, this activity can be determined using  
CC fluorescence or absorbance readouts. The present invention describes a  
CC method that identifies modulators of the target protein, which is a  
CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiac,  
CC immunomodulators and antiinflammatory. Accordingly, through gene  
CC therapy, they can be used for the treatment of cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
CC polynucleotide sequence is human KID DNA (Seqid 7) encoding a full length  
CC KID enzyme of the invention.

SQ Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

Query Match 100.0%; Score 1464; DB 10; Length 1464;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTA 60  
DB 1 ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGCGTA 60

QY 61 AGGGTGGCTGTGCGACTGCGGCAATTTGTGGATGGAAACAGCGGGAGCAAGTATCCCCC 120  
DB 61 AGGGTGGCTGTGCGACTGCGGCAATTTGTGGATGGAAACAGCGGGAGCAAGTATCCCCC 120

QY 121 TGTGTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180  
DB 121 TGTGTGGGGGATGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAG 180

QY 181 ACTCTCAATACACAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGGACATCTAT 240  
DB 181 ACTCTCAATACACAGTTTGTATGCTTCTATGGGGAGAGAGTACTCAGCAGGACATCTAT 240

QY 241 GCAGGTTCACTGAGCCCATCTTAAGCACTTCTGTAAGGGGAGATGCCAGTGTGCTT 300  
DB 241 GCAGGTTCACTGAGCCCATCTTAAGCACTTCTGTAAGGGGAGATGCCAGTGTGCTT 300

QY 301 GCCTATGGACCCACAGAGCTGGGAACGACACACATGCTGGGGAGCCCGAGCAACCT 360  
DB 301 GCCTATGGACCCACAGAGCTGGGAACGACACACATGCTGGGGAGCCCGAGCAACCT 360

QY 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCATGGACCTCTCATGGAGGGAGGGTGCAG 420  
DB 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCATGGACCTCTCATGGAGGGAGGGTGCAG 420

QY 421 GGCGGGCCATGGGGCCCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480  
DB 421 GGCGGGCCATGGGGCCCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480

QY 481 TTAGACTCTCTGGACCTCTGTCGGAGACTGCTTAATCCGAGAACTGCCGGGGAAAT 540  
DB 481 TTAGACTCTCTGGACCTCTGTCGGAGACTGCTTAATCCGAGAACTGCCGGGGAAAT 540

QY 541 ATCTGATTCGGGTCTCTCCAGAACCCATAGTAGTCTTGTGATTTTGGAGCGGCAC 600  
DB 541 ATCTGATTCGGGTCTCTCCAGAACCCATAGTAGTCTTGTGATTTTGGAGCGGCAC 600

QY 601 TTCTGTCAGGACCTCGAAATCGACTGTAGGACCCACCGGCTCAACAGAGCGTCTCTCC 660  
DB 601 TTCTGTCAGGACCTCGAAATCGACTGTAGGACCCACCGGCTCAACAGAGCGTCTCTCC 660

QY 661 CGCAGTCTATGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACTTTGGCCCCATTTGCG 720  
DB 661 CGCAGTCTATGCTGCTCTCTGCTCAAGGTGGACAGCGGGAACTTTGGCCCCATTTGCG 720

QY 721 CAGCGAGAGGAAACTCTACTGATTGACTTGGCTGGGTGAGGACCAACCGCGGCACA 780  
DB 721 CAGCGAGAGGAAACTCTACTGATTGACTTGGCTGGGTGAGGACCAACCGCGGCACA 780

QY 781 GGCACACAGGCGCTTCGGCTAAAGAGAGTGGAGCCCATCAACACCTCCCTGTTGTCTG 840  
DB 781 GGCACACAGGCGCTTCGGCTAAAGAGAGTGGAGCCCATCAACACCTCCCTGTTGTCTG 840

QY 841 GGCACAGTGGTAGATGGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGGACAGCAAG 900  
DB 841 GGCACAGTGGTAGATGGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGGACAGCAAG 900

QY 901 CTCACTCCCTTATTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAAC 960  
DB 901 CTCACTCCCTTATTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTCGCAAC 960

QY 961 ATTGCCCCCTGAGAGAGCGTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020  
DB 961 ATTGCCCCCTGAGAGAGCGTTCTACTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020

QY 1021 TCCAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTTCAGGCTTCATGCTTGGGA 1080  
DB 1021 TCCAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTTCAGGCTTCATGCTTGGGA 1080

QY 1081 CCTGTTAAGCTGTCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140  
DB 1081 CCTGTTAAGCTGTCTCAGAAAGATTGCTTGGTCCACAGAGGCAAGAGAGCCGAGGC 1140

QY 1141 CCTGAGGAGAGGAGATTGGGAGCCCTGAGGCCCATGGCAGCTTCCAGCCTTCGCTCCCAG 1200  
DB 1141 CCTGAGGAGAGGAGATTGGGAGCCCTGAGGCCCATGGCAGCTTCCAGCCTTCGCTCCCAG 1200

QY 1201 AAATCAGCCCCCTTACAGAGCTAAGAGCATGGACCCGGCCATGCTGGAGGCGCTCTC 1260  
DB 1201 AAATCAGCCCCCTTACAGAGCTAAGAGCATGGACCCGGCCATGCTGGAGGCGCTCTC 1260

QY 1261 AGCTTGGACCGCTCTGCTTCCCTCCAGGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCA 1320  
DB 1261 AGCTTGGACCGCTCTGCTTCCCTCCAGGGAGCCAGGGGCGCCCTCTGTTGAGTACCCCA 1320

QY 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGG 1380  
DB 1321 AAGCGAGAGCGGATGCTGCTAATGAAGACAGTGAAGAGAGGACCTAGAGATTGAGAGG 1380

QY 1381 CTTAAGACGAGCAAAAGAACTGGAGCCCAAGATTTGGCCCAAGAGGCTGAGGAAAAG 1440  
DB 1381 CTTAAGACGAGCAAAAGAACTGGAGCCCAAGATTTGGCCCAAGAGGCTGAGGAAAAG 1440

QY 1441 GAGAACCATTTGCCACATGTGA 1464  
DB 1441 GAGAACCATTTGCCACATGTGA 1464

RESULT 2  
ADQ60233  
ID ADQ60233 standard; DNA; 1464 BP.  
XX  
AC ADQ60233;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Human microtubule motor protein DNA #4.  
XX  
KW Human; microtubule motor protein; gene; ds;  
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
KW cardiac hypertrophy; immune disorder; inflammation;  
KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;  
KW graft rejection; inflammatory bowel disease; angioplasty.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 1. .1464  
FT /\*tag= a  
FT /product= "Human microtubule motor protein #4"



QY	1441	GAGAACCAATGTCACCAATGGA	1464
Db	1441	GAGAACCAATGTCACCAATGGA	1464
RESULT 3			
AD23339	ID		
AD23339	standard; DNA; 1538 BP.		
XX	AD23339;		
AC	AD23339;		
XX	18-DEC-2003 (first entry)		
DE	DNA encoding the human kinesin-like DNA binding protein (KID) (seqID 3).		
XX	human; gene; db; motor domain; kinesin-like DNA binding protein; KID;		
KW	cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;		
KW	cancer; hyperplasia; restenosis; cellular proliferation disorder;		
KW	cardiac hypertrophy; immune disorder; inflammation.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	1..1538		
FT	/*tag= a		
FT	/product= "KID protein"		
FT	/transl_except= (pos: 4..5; aa: Pro)		
FT	/note= "This codon has an apparent 1 nucleotide deletion		
FT	that alters the reading frame"		
XX	XX		
PN	US6387644-B1.		
XX	XX		
PD	14-MAY-2002.		
XX	XX		
PP	28-NOV-2000; 2000US-00724224.		
XX	XX		
PR	20-APR-1999; 99US-00295612.		
PR	20-JUN-2000; 2000US-00597292.		
XX	XX		
PA	(CYTO-) CYTOKINETICS INC.		
XX	XX		
PI	Beraud C;		
XX	XX		
DR	WPI; 2003-706919/67.		
DR	P-PSDB; ADC23340.		
XX	XX		
PT	Identifying a candidate agent as modulator of function of a target		
PT	protein for treating cellular proliferation disorders by adding a		
PT	candidate agent to a mixture of the target protein that		
PT	directly/indirectly produces ADP or phosphate.		
XX	XX		
PS	Disclosure; SEQ ID NO 3; 26pp; English.		
XX	XX		
CC	This invention relates to a novel method for high throughput screening		
CC	systems used to identify compounds for the treatment of cellular		
CC	proliferation disorders. Specifically, it refers to candidate agents that		
CC	are capable of modulating the activity of target proteins having motor		
CC	domains, such that the target protein directly or indirectly produces ADP		
CC	or phosphate. Furthermore, this activity can be determined using		
CC	fluorescence or absorbance readouts. The present invention describes a		
CC	method that identifies modulators of the target protein, which is a		
CC	kinesin-like DNA binding protein (known as KID) as cytostatic, cardiants,		
CC	immunomodulators and antiinflammatories. Accordingly, through gene		
CC	therapy, they can be used for the treatment of cancer, hyperplasias,		
CC	restenosis, cardiac hypertrophy, immune disorders and inflammation. This		
CC	polynucleotide sequence is human KID DNA (SeqID 3) encoding a full length		
CC	KID enzyme of the invention.		
XX	XX		
SQ	Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;		
Query Match 99.8%; Score 1461; DB 10; Length 1538;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

QY	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGCGTAAGG	63
Db	78	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCCACCTCCAGCTCGCGTAAGG	137
QY	64	GTGGCTGTGGGACTGCGGCCCATTTGTGGATGGAAACAGCGGAGAGCAAGTATCCCCCTGT	123
Db	138	GTGGCTGTGGGACTGCGGCCCATTTGTGGATGGAAACAGCGGAGAGCAAGTATCCCCCTGT	197
QY	124	GTGGGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCCAGAGAGACT	183
Db	198	GTGGGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCCAGAGAGACT	257
QY	184	CTCAAAATACCAAGTTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
Db	258	CTCAAAATACCAAGTTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
QY	244	GGTTCAGTGTCAGGCCCATCTTAAGGCATTTGCTGAAGGGCAGAAATGCCAGTGTGCTGCC	303
Db	318	GGTTCAGTGTCAGGCCCATCTTAAGGCATTTGCTGAAGGGCAGAAATGCCAGTGTGCTGCC	377
QY	304	TATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCACAGAGCAACTGGG	363
Db	378	TATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCACAGAGCAACTGGG	437
QY	364	GTGATCCCGCGGCTCTCATGGACCTCTCTGAGCTCACAAGGGAGGGTGCAGGGGC	423
Db	438	GTGATCCCGCGGCTCTCATGGACCTCTCTGAGCTCACAAGGGAGGGTGCAGGGGC	497
QY	424	CGGCCATGGGCGCTTTCTGTCACATGCTTACCTAGAGATCTACCAGAGAGGATATTA	483
Db	498	CGGCCATGGGCGCTTTCTGTCACATGCTTACCTAGAGATCTACCAGAGAGGATATTA	557
QY	484	GACCTCTGGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAATATC	543
Db	558	GACCTCTGGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAATATC	617
QY	544	CTGATTCGGGCTCTCTCCAGAAAGCCCATCATAGTAGCTTTGCTGATTTTGAGGGGCACTTC	603
Db	618	CTGATTCGGGCTCTCTCCAGAAAGCCCATCATAGTAGCTTTGCTGATTTTGAGGGGCACTTC	677
QY	604	CTGCCAGCCAGTCGAATCGGACTCTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	663
Db	678	CTGCCAGCCAGTCGAATCGGACTCTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	737
QY	664	AGTCATGCTGCTCTCTGGTCAAGGTGGACCGGGAACGTTTGCCCCCAATTTGCCAG	723
Db	738	AGTCATGCTGCTCTCTGGTCAAGGTGGACCGGGAACGTTTGCCCCCAATTTGCCAG	797
QY	724	CGAGAGGGGAAAACTCTTACCTGATTGATTTGGTCTGAGGACACCCGGGCGACAGGC	783
Db	798	CGAGAGGGGAAAACTCTTACCTGATTGATTTGGTCTGAGGACACCCGGGCGACAGGC	857
QY	784	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCGGGC	843
Db	858	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCGGGC	917
QY	844	AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTTATCGGACAGCAAGCTC	903
Db	918	AAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTCGTGTACCTTTATCGGACAGCAAGCTC	977
QY	904	ACTCGCTATTGTCAGAGCTCTCTGGTGGCTGAGCCACAGTATCTTTATTCGCAACATT	963
Db	978	ACTCGCTATTGTCAGAGCTCTCTGGTGGCTGAGCCACAGTATCTTTATTCGCAACATT	1037
QY	964	GCCCTCTGAGAGCGCTTCTACCTTAGACACAGTCTCCGCACTCAACTTTGTGCGAGGTCC	1023
Db	1038	GCCCTCTGAGAGCGCTTCTACCTTAGACACAGTCTCCGCACTCAACTTTGTGCGAGGTCC	1097
QY	1024	AAGGAGGTGATCAATCGGCTTTTATCAATGAGAGCTGAGGCTCATGCTTGGGACCT	1083
Db	1098	AAGGAGGTGATCAATCGGCTTTTATCAATGAGAGCTGAGGCTCATGCTTGGGACCT	1157





QY 724 CGAGAGGGAAGAACTCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGGCCACAGC 783  
DB |||||||  
QY 798 CGAGAGGGAAGAACTCTACCTGATTGACTTGGCTGGGTGAGAGCAACCGGGCCACAGC 857  
DB |||||||  
QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTCTGTTGTCTCTGGCC 843  
DB |||||||  
QY 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTCTGTTGTCTCTGGCC 917  
DB |||||||  
QY 844 AAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903  
DB |||||||  
QY 918 AAGTGGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 977  
DB |||||||  
QY 904 ACTCGCTATTGAGAGACTCTCTGGTGGCTCAGGCCACAGTATCCTTATTGCGCAACTT 963  
DB |||||||  
QY 978 ACTCGCTATTGAGAGACTCTCTGGTGGCTCAGGCCACAGTATCCTTATTGCGCAACTT 1037  
DB |||||||  
QY 964 GCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCGCACTCAACTTGTCTCGAGGCTC 1023  
DB |||||||  
QY 1038 GCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCGCACTCAACTTGTCTCGAGGCTC 1097  
DB |||||||  
QY 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1083  
DB |||||||  
QY 1098 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAGCTCATGCTTGGGACCT 1157  
DB |||||||  
QY 1084 GTTAGCTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1143  
DB |||||||  
QY 1158 GTTAGCTCTCAGAAAGAAATGCTTGTGTCACAGAGGCAAGAGAGCCCGAGGCCCT 1217  
DB |||||||  
QY 1144 GAGGAGAGGAGATTGGGAGCCCTGAGCCCATGCGAGCTTCCAGCTTCTCCCTCCAGAAA 1203  
DB |||||||  
QY 1218 GAGGAGAGGAGATTGGGAGCCCTGAGCCCATGCGAGCTTCCAGCTTCTCCCTCCAGAAA 1277  
DB |||||||  
QY 1204 CTCAGCCCCCTCAGAAAGCTAAGCAGCATGAGCCCGCCCATGCTGAGGCGCTCTCTCAGC 1263  
DB |||||||  
QY 1278 CTCAGCCCCCTCAGAAAGCTAAGCAGCATGAGCCCGCCCATGCTGAGGCGCTCTCTCAGC 1337  
DB |||||||  
QY 1264 TTGAGCGCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323  
DB |||||||  
QY 1338 TTGAGCGCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397  
DB |||||||  
QY 1324 CGAGAGCGGATGCTGCTTAAGACAGTAGAGAGAGGACCTAGAGATTGAGAGGCTT 1383  
DB |||||||  
QY 1398 CGAGAGCGGATGCTGCTTAAGACAGTAGAGAGAGGACCTAGAGATTGAGAGGCTT 1457  
DB |||||||  
QY 1384 AAGACGAAGCAAAAGAACTGGAGGCCAGATGTTGGCCAGAGCTGAGGAAAGGAG 1443  
DB |||||||  
QY 1458 AAGACGAAGCAAAAGAACTGGAGGCCAGATGTTGGCCAGAGGCTGAGGAAAGGAG 1517  
DB |||||||  
QY 1444 AACCATGTCCTCAATGTGA 1464  
DB |||||||  
QY 1518 AACCATGTCCTCAATGTGA 1538  
DB |||||||

RESULT 5  
ADQ09241  
ID ADQ09241 standard; cDNA; 2097 BP.  
XX AC ADQ09241;  
XX AC ADQ09241;  
DT 23-SEP-2004 (first entry)  
XX Human KNSL4 encoding cDNA SEQ ID NO:426.  
XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;  
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;  
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;  
KW angiogenic; antiinflammatory; cardiovascular; cytostatic;  
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;  
KW human; gene; ss.  
XX Homo sapiens.  
OS  
XX W02004055050-A2.  
PN

XX 01-JUL-2004.  
PD 10-DEC-2003; 2003WO-IB006434.  
XX 10-DEC-2002; 2002US-0432699P.  
XX 03-JUL-2003; 2003US-0485027P.  
XX (ENDO-) ENDOCUBE SAS.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
PI Girard J, Amalric F, Roussigne M, Clouaire T;  
XX WPI; 2004-525034/50.  
DR P-PSDB; ADQ09240.  
XX  
PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)  
PT responsive gene for preventing or treating e.g. cancer or inflammation,  
PT comprises modulating the interaction of a THAP polypeptide with a nucleic  
PT acid.  
XX  
PS Example 47; SEQ ID NO 426; 612pp; English.  
XX  
CC The present invention describes a method for modulating the expression of  
CC a thanatos (death)-associated protein (THAP) responsive gene. The method  
CC comprises modulating the interaction of a THAP-family polypeptide or its  
CC biological fragment with a nucleic acid, and so enhancing or repressing  
CC the expression of the THAP responsive gene. Also described: (1) a method  
CC of modulating the expression of a gene responsive to a THAP/chemokine  
CC complex; (2) a pharmaceutical composition comprising a THAP responsive  
CC element in a pharmaceutical carrier; (3) a transcription factor decoy  
CC consisting essentially of a THAP responsive element; (4) a cell  
CC comprising a transcription factor decoy described above; (5) methods of  
CC modulating the interaction between a nucleic acid and a THAP-family  
CC polypeptide or its biological fragment, or a nucleic acid and a  
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a  
CC cell comprising a viral vector which comprises a promoter operably linked  
CC to a nucleic acid encoding a THAP-family polypeptide or its biological  
CC fragment; (7) a method of constructing a cell which expresses a  
CC recombinant THAP-family polypeptide; (8) a method of ameliorating  
CC symptoms associated with a condition mediated by a THAP/chemokine complex  
CC ; (9) methods of identifying a test compound that modulates transcription  
CC at a THAP responsive element or that modulates the transport of a  
CC chemokine into the nucleus; (10) methods for reducing the symptoms  
CC associated with a condition selected from excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative  
CC diseases; symptoms associated with a condition resulting from the  
CC activity of a chemokine or a THAP-family polypeptide in an individual; or  
CC symptoms associated with transcriptional repression or activation  
CC mediated by a THAP-family polypeptide in an individual; (11) a vector  
CC comprising a THAP responsive promoter operably linked to a nucleic acid  
CC encoding a detectable product; (12) a genetically engineered cell  
CC comprising the vector described above or that expresses a THAP-family  
CC polypeptide or its biological fragment; (13) an in vitro transcription  
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,  
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-  
CC family polypeptide that does not bind to a chemokine. The pharmaceutical  
CC composition has antiangiogenic, antiinflammatory, cardiovascular,  
CC cytoskeletal, neuroprotective and osteopathic activities, and can be used  
CC as a THAP and THAP synthesis modulator. The composition can be used for  
CC modulating the expression of a THAP responsive gene. Modulation is useful  
CC for reducing symptoms of conditions such as excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative  
CC diseases. The present sequence is used in the exemplification of the  
CC present invention.  
SQ Sequence 2097 BP; 483 A; 592 C; 628 G; 394 T; 0 U; 0 Other;

Query Match 99.6%; Score 1458; DB 12; Length 2097;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	4	G	G	T	C	G	T	G	T	C	G	G	C	T	A	A	G	A	G	T	T	G	G	A	G	T	T	G	G	A	G	T	T	G	T	C	G	T	C	A	C	T	C	A	G	T	C	G	C	T	A	A	G	63
Db	100	G	G	T	C	G	T	G	T	C	G	G	C	T	A	A	G	A	G	T	T	G	G	A	G	T	T	G	G	A	G	T	T	G	T	C	G	T	C	A	C	T	C	A	G	T	C	G	C	T	A	A	G	159
Qy	64	G	T	G	G	T	G	T	G	C	G	A	C	T	T	T	G	T	G	A	T	G	G	A	C	A	G	C	G	G	A	G	A	G	T	G	T	C	C	C	C	C	C	T	G	T	123							
Db	160	G	T	G	G	T	G	T	G	C	A	C	T	T	T	T	G	T	G	A	T	G	G	A	C	A	G	C	G	G	A	G	A	G	T	G	T	C	C	C	C	C	C	T	G	T	219							
Qy	124	G	T	G	G	G	G	C	A	T	G	A	C	T	G	T	C	T	A	G	A	T	T	G	T	A	A	C	T	G	A	G	A	A	C	C	A	G	A	G	A	C	T	A	T	G	C	A	183					
Db	220	G	T	G	G	G	G	C	A	T	G	A	C	T	G	T	C	T	A	G	A	T	T	G	T	A	A	C	T	G	A	G	A	A	C	C	A	G	A	G	A	C	T	A	T	G	C	A	279					
Qy	184	C	T	C	A	A	T	A	C	C	A	G	T	T	G	A	T	G	C	T	T	A	T	G	G	G	A	G	A	G	A	G	A	G	A	G	A	C	A	T	A	T	G	C	A	243								
Db	280	C	T	C	A	A	T	A	C	C	A	G	T	T	T	G	A	T	G	C	T	T	A	T	G	G	G	A	G	A	G	A	G	A	G	A	C	A	T	A	T	G	C	A	339									
Qy	244	G	G	T	C	A	G	T	C	A	G	C	C	A	T	C	T	A	A	G	C	A	T	T	G	T	G	G	A	G	G	C	A	G	A	T	G	C	C	A	G	A	T	G	T	G	C	303						
Db	340	G	G	T	C	A	G	T	C	A	G	C	C	A	T	C	T	A	A	G	C	A	T	T	G	T	G	G	A	G	G	C	A	G	A	T	G	C	C	A	G	A	T	G	T	G	C	399						
Qy	304	T	A	T	G	A	C	C	A	C	A	G	A	G	T	G	G	A	A	G	A	C	A	C	A	A	T	G	T	G	G	A	G	C	C	A	G	A	C	A	A	C	T	G	G	363								
Db	400	T	A	T	G	A	C	C	A	C	A	G	A	G	T	G	G	A	A	G	A	C	A	C	A	A	T	G	T	G	G	A	G	C	C	A	G	A	C	A	A	C	T	G	G	459								
Qy	364	G	T	G	A	T	C	C	G	C	G	G	G	T	C	T	C	A	T	G	A	C	T	C	T	C	A	A	G	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	423							
Db	460	G	T	G	A	T	C	C	G	C	G	G	G	T	C	T	C	A	T	G	A	C	T	C	T	C	T	C	A	A	G	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	519							
Qy	424	C	G	G	C	A	T	G	G	C	C	T	T	C	T	G	T	C	A	C	A	T	G	T	T	A	C	T	T	A	G	A	G	A	T	C	A	C	A	G	A	G	A	G	A	G	483							
Db	520	C	G	G	C	A	T	G	G	C	C	T	T	C	T	G	T	C	A	C	A	T	G	T	T	A	C	T	T	A	G	A	G	A	T	C	A	C	A	G	A	G	A	G	A	G	579							

Qy	1084	GTTHAGCTGTCTCAGAAAGAATTCCTTGGTCCACCCAGAGGCAAGAGAGCCCGAGGCGCCT	1143
Db	1180	GTTTAAGCTGTCTCAGAAAGAATTCCTTGGTCCACCCAGAGGCAAGAGAGCCCGAGGCGCCT	1239
Qy	1144	GAGGAAGAGAGATTGGGAGCCCTTGAGCCCATGCGAGCTCCAGGCTCTGGCTCTCCAGAAA	1203
Db	1240	GAGGAAGAGAGATTGGGAGCCCTTGAGCCCATGCGAGCTCCAGGCTCTGGCTCTCCAGAAA	1299
Qy	1204	CTCAGCCCCCTACAGAAAGCTTAAGCAGCATGGACCCCGGCCATGCTGGAGCGCCTCCTCAGC	1263
Db	1300	CTCAGCCCCCTACAGAAAGCTTAAGCAGCATGGACCCCGGCCATGCTGGAGCGCCTCCTCAGC	1359
Qy	1264	TTGACCGTCTGCTTGCCCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1323
Db	1360	TTGACCGTCTGCTTGCCCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG	1419
Qy	1324	CGAGGCGGATGGTGCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT	1383
Db	1420	CGAGGCGGATGGTGCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT	1479
Qy	1384	AAGACGAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGGCTGAGGAAAAAGGAG	1443
Db	1480	AAGACGAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGGCTGAGGAAAAAGGAG	1539
Qy	1444	AACCATGTGCCCAATG	1461
Db	1540	AACCATGTGCCCAATG	1557
RESULT 6			
ADK40995			
ID	ADK40995 standard; DNA; 2099 BP.		
XX			
AC	ADK40995;		
XX			
DT	06-MAY-2004 (first entry)		

RESULT 6  
ADK40995

ID ADK40995 standard; DNA; 2099 BP.

AC ADK40995;

DT 06-MAY-2004 (first entry)

DE Novel human kinase gene #15.

cytostatic; immunomodulator; cardiant; neuroprotective; nootropic;  
KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
KW cancer; peripheral nervous system; central nervous system;  
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
KW amyotrophic lateral sclerosis; viral infection; prion infection;  
KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
KW attention disorder; cognition disorder; hypotension; hypertension;  
KW psychotic disorder; neurological disorder; dyskinesia;  
KW metabolic disorder; organ transplant rejection; enzyme; gene; ds.

OS Homo sapiens.

AA  
PN  
WO2003057841-A2.

17-JUL-2003

31-DEC-2002: 2002WO-US041687.

31-DEC-2001: 2001US-0343169P.

XX PA (GRIG/) GRIGORIEV I V.

XX  
PA (SUDA/) SUDARSANAM S.

PI Grigoriev IV, Sudarsanam S;  
yy

DR WPI; 2003-587115/55.

PT New isolated, enriched or purified nucleic acid molecule encoding a  
PT kinase polypeptide, useful for treating cancer, immune-related diseases,  
PT cardiovascular disease, brain or neuronal-associated diseases and  
PT metabolic disorders.

PS Claim 33; SEQ ID NO 102; 491pp: English.

•



KW wound; infectious disease; ss.  
 OS Homo sapiens.  
 XX WO200055174-A1.  
 XX 21-SEP-2000.  
 XX 08-MAR-2000; 2000WO-US005988.  
 XX 12-MAR-1999; 99US-0124270P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX (ROSE/) ROSEN C A.  
 XX Rosen CA, Ruben SM;  
 XX WPI: 2000-587513/55.  
 XX P-PSDB; AAB56650.  
 XX  
 XX Prostate cancer associated gene sequences, referred to as prostate cancer  
 XX antigens, useful for treatment, prevention, and diagnosis of disorders  
 XX such as prostate cancer.  
 XX  
 XX Claim 1; Page 805-806; 2338pp; English.  
 XX  
 XX AAF15566 to AAF16505 encode the human prostate cancer associated  
 XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 XX The prostate cancer antigens can have neuroprotective, cytostatic,  
 XX cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
 XX nephrotropic, anti-infective, gynecological and antibacterial activities,  
 XX and can be used in gene therapy. The prostate cancer antigen  
 XX polynucleotides may be used for detection of prostate cancer, chromosome  
 XX identification, as chromosome markers, and for numerous other diagnostic  
 XX or research purposes. The prostate cancer antigens may be used to treat  
 XX disorders such as neural, immune, muscular, reproductive,  
 XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
 XX AAB57303 represent sequences used in the exemplification of the present  
 XX invention  
 XX  
 XX Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;  
 XX  
 XX Query Match 98.6%; Score 1443.2; DB 3; Length 2104;  
 XX Best Local Similarity 99.4%; Pred. No. 0;  
 XX Matches 1450; Conservative 7; Mismatches 1; Indels 1; Gaps 1;  
 XX  
 XX 4 GGTGGCTGTGGCTTAAAGAGATGGAGCTACTGCTGTCCACCTCCAGCTCGGTAAGG 63  
 XX 67 GGTGGCTGTGGCTTAAAGAGATGGAGCTACTGCTGTCCACCTCCAGCTCGGTAAGG 126  
 XX  
 XX 64 GTGGCTGTGGCTTAAAGAGATGGAGCTACTGCTGTCCACCTCCAGCTCGGTAAGG 123  
 XX 127 GTGGCTGTGGCTTAAAGAGATGGAGCTACTGCTGTCCACCTCCAGCTCGGTAAGG 186  
 XX  
 XX 124 GTGGGGGATGGAGCTGCTCTAGAGATGCTAACTGGAGAACACACAGGAGACT 183  
 XX 187 GTGGGGGATGGAGCTGCTCTAGAGATGCTAACTGGAGAACACACAGGAGACT 246  
 XX  
 XX 184 CTCMAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTAGCAGGACATCTATGCA 243  
 XX 247 CTCMAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTAGCAGGACATCTATGCA 306  
 XX  
 XX 244 GGTTCAGTCAGCCCATCTTAAGCACTTCTGAGGGGAGAGTCCAGTGTGCTTGGC 303  
 XX 307 GGTTCAGTCAGCCCATCTTAAGCACTTCTGAGGGGAGAGTCCAGTGTGCTTGGC 366  
 XX  
 XX 304 TATGGAGCCACAGAGCTGGGAGAGCGCACCAATGCTGGGAGCCCGAGAGCAACCTGGG 363  
 XX 367 TATGGAGCCACAGAGCTGGGAGAGCGCACCAATGCTGGGAGCCCGAGAGCAACCTGGG 426  
 XX  
 XX 364 GTGATCCCGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAGGTCCGAGGGC 423

Db 427 GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGGAGGCTCCGAGGGC 486  
 Qy 424 CGGCATGGGGCTCTTCTGTCAACCATGCTTACCTAGAGATCTACACAGAGAGGTATTA 483  
 Db 487 CGGCATGGGNCCTTCTGTCAACCATGCTTACCTAGAGATCTACACAGAGAGGTATTA 546  
 Qy 484 GACCTCTGGACCTGCTTCCGAGAGACCTGGTAATCCGAGAGAGTTCGCGGGGGAATATC 543  
 Db 547 GACCTCTGGACCTGCTTCCGAGAGACCTGGTAATCCGAGAGAGTTCGCGGGGGAATATC 606  
 Qy 544 CTGATTCGGGGTCTCTCCAGAGACCTCATAGTAGCTTGTGATTTTGGAGGGGACATTC 603  
 Db 607 CTGATTCGGGGTCTCTCCAGAGACCTCATAGTAGCTTGTGATTTTGGAGGGGACATTC 666  
 Qy 604 CTGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCTCTCTCCGC 663  
 Db 667 CTGCAGCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACACAGCGCTCTCTCCGC 726  
 Qy 664 AGTCATGCTGTCTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 723  
 Db 727 AGTCATGCTGTCTCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTC 786  
 Qy 724 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTCGAGGAGCAACCGGCGCACAGGC 783  
 Db 787 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTCGAGGAGCAACCGGCGCACAGGC 846  
 Qy 784 AACAGGGCTTCCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGCGGC 843  
 Db 847 AACAGGGCTTCCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGCGGC 906  
 Qy 844 AAAGTGTAGATGCGCTGTAATCAGGGGCTCTCTGTCGTACCTTATCGGAGCAGCAAGCTC 903  
 Db 907 AAAGTGTAGATGCGCTGTAATCAGGGGCTCTCTGTCGTACCTTATCGGAGCAGCAAGCTC 966  
 Qy 904 ACTGCGCTATTCAGGACTCTCTGGTGGCTCAGCCACACAGTATCTTTATGCGCAACAT 963  
 Db 967 ACTGCGCTATTCAGGACTCTCTGGTGGCTCAGCCACACAGTATCTTTATGCGCAACAT 1026  
 Qy 964 GCCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTGTGCGCAGGTC 1023  
 Db 1027 GCCCTCAGAGAGCTTCTACCTAGACACAGTCTCCGACCTCAACTTGTGCGCAGGTC 1086  
 Qy 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCCCTCATGCTTGGGAGCT 1083  
 Db 1087 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGAGCCCTCATGCTTGGGAGCT 1146  
 Qy 1084 GTTAAGCTGTCTCAGAAAGAAATTTGCTTGGTCCACAGAGGCAAGAGAGCCGAGGCGCT 1143  
 Db 1147 GTTAAGCTGTCTCAGAAAGAAATTTGCTTGGTCCACAGAGGCAAGAGAGCCGAGGCGCT 1206  
 Qy 1144 GAGGAGAGGAGATTTGGGAGCCCTGAGCCCATGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1203  
 Db 1207 GAGGAGAGGAGATTTGGGAGCCCTGAGCCCATGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1266  
 Qy 1204 CTCAGGCCCCCTACAGAGCTTAAGCAGCATGAGACCCCGCCCATGCTTGGAGCGCTCTCT 1262  
 Db 1267 CTCAGGCCCCCTACAGAGCTTAAGCAGCATGAGACCCCGCCCATGCTTGGAGCGCTCTCT 1326  
 Qy 1263 CTTGGAGCCGCTGCTTGGCTTCCAGAGGAGCCAGGGGGGCGCCCTCTGTTGAGTACCCAAA 1322  
 Db 1327 CTTGGAGCCGCTGCTTGGCTTCCAGAGGAGCCAGGGGGGCGCCCTCTGTTGAGTACCCAAA 1386  
 Qy 1323 GCAGAGCGGATGCTCTAATGAAGACAGTAGAGAGAGAGGAGGAGGAGGAGGAGGAGGCT 1382  
 Db 1387 GCAGAGCGGATGCTCTAATGAAGACAGTAGAGAGAGAGGAGGAGGAGGAGGAGGAGGCT 1446  
 Qy 1383 TAAGAGCAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCGAGAGGCTGAGGAGGAGGAG 1442  
 Db 1447 TAAGAGCAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCGAGAGGCTGAGGAGGAGGAG 1506  
 Qy 1443 GAACCAATTTGCCCAATG 1461  
 Db 1507 GAACCAATTTGCCCAATG 1525

RESULT 8  
 ADC23341  
 ID ADC23341 standard; DNA; 1041 BP.  
 XX  
 AC ADC23341;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 5).  
 XX  
 KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
 KW cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;  
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
 KW cardiac hypertrophy; immune disorder; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1041  
 FT /\*tag= a  
 FT /product= "KID protein"  
 XX  
 PN US6387644-B1.  
 XX  
 PD 14-MAY-2002.  
 XX  
 PF 28-NOV-2000; 2000US-00724224.  
 XX  
 PR 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 XX  
 PA (CYTO-) CYTOKINETICS INC.  
 XX  
 PI Beraud C;  
 XX  
 DR WPI; 2003-706919/57.  
 XX  
 PP P-PSDB; ADC23342.  
 XX  
 PT Identifying a candidate agent as modulator of function of a target  
 PT protein for treating cellular proliferation disorders by adding a  
 PT candidate agent to a mixture of the target protein that  
 PT directly/indirectly produces ADP or phosphate.  
 XX  
 PS Disclosure; SEQ ID NO 5; 26pp; English.  
 XX  
 CC This invention relates to a novel method for high throughput screening  
 CC systems used to identify compounds for the treatment of cellular  
 CC proliferation disorders. Specifically, it refers to candidate agents that  
 CC are capable of modulating the activity of target proteins having motor  
 CC domains, such that the target protein directly or indirectly produces ADP  
 CC or phosphate. Furthermore, this activity can be determined using  
 CC fluorescence or absorbance readouts. The present invention describes a  
 CC method that identifies modulators of the target protein, which is a  
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiants,  
 CC immunomodulators and antiinflammatory. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polynucleotide sequence is human KID DNA (SeqID 5) encoding a full length  
 CC KID enzyme of the invention.  
 XX  
 SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;  
 Query Match 70.9%; Score 1038.4; DB 10; Length 1041;  
 Best Local Similarity 99.9%; Pred. No. 3.le-278;  
 Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGGGTCGCTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCCGGTA 60  
 DB 1 ATGGGTCGCTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCCGGTA 60  
 QY 61 AGGGTGGCTGTGGCACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTGATCCCCC 120

DB 61 AGGGTGGCTGTGGCACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTGATCCCCC 120  
 QY 121 TGTGTGCGGGGCATGGAACAGCTGTCTCTAGAGATTGCTAACTGGAGAAACACAGGAG 180  
 DB 121 TGTGTGCGGGGCATGGAACAGCTGTCTCTAGAGATTGCTAACTGGAGAAACACAGGAG 180  
 QY 181 ACTCTCAATACCAAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT 240  
 DB 181 ACTCTCAATACCAAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT 240  
 QY 241 GCAGGTTTCAGTGCAGCCCATCTTAAAGGCATTTGCTGGAAGGCGCAAGATGCCAGTGTCT 300  
 DB 241 GCAGGTTTCAGTGCAGCCCATCTTAAAGGCATTTGCTGGAAGGCGCAAGATGCCAGTGTCT 300  
 QY 301 GCCTATGAGCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGAGCAAGCAACT 360  
 DB 301 GCCTATGAGCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGGAGCAAGCAACT 360  
 QY 361 GGGGTGATCCCGCGGGCTCTCATGACCTCTGTCAGCTCACAAAGGAGAGAGGTGCCGAG 420  
 DB 361 GGGGTGATCCCGCGGGCTCTCATGACCTCTGTCAGCTCACAAAGGAGAGAGGTGCCGAG 420  
 QY 421 GGC CGGCCCATGGGCCCTTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTA 480  
 DB 421 GGC CGGCCCATGGGCCCTTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGTA 480  
 QY 481 TTAGACCTCTGAGACCTGCTTCGGGAGACCTGTAATCCGAGAGAGACTGCCGGGGGAT 540  
 DB 481 TTAGACCTCTGAGACCTGCTTCGGGAGACCTGTAATCCGAGAGAGACTGCCGGGGGAT 540  
 QY 541 ATCTCTGATTCCGGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTTGTGATTTGAGCGGCAC 600  
 DB 541 ATCTCTGATTCCGGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTTGTGATTTTTCAGCGGCAC 600  
 QY 601 TTCTTGCAGCCAGTCGAAATCGGACTGTAGAGGCCACCGGCTCAACAGGCTCTCTCC 660  
 DB 601 TTCTTGCAGCCAGTCGAAATCGGACTGTAGAGGCCACCGGCTCAACAGGCTCTCTCC 660  
 QY 661 CGCAGTCATGCTGTCTCTGCTGAGTGGACCGGGAACGTTTGGCCCCCATTTGCTGC 720  
 DB 661 CGCAGTCATGCTGTCTCTGCTGAGTGGACCGGGAACGTTTGGCCCCCATTTGCTGC 720  
 QY 721 CAGCGAGAGGGGAAAACCTTACCTGATTGACTTTGGCTGGGTCTAGAGCAACCGGCGACA 780  
 DB 721 CAGCGAGAGGGGAAAACCTTACCTGATTGACTTTGGCTGGGTCTAGAGCAACCGGCGACA 780  
 QY 781 GGCACAAAGGGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCTG 840  
 DB 781 GGCACAAAGGGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCTG 840  
 QY 841 GGCACAAAGTGTAGATGCGCTGAAATCAGGCGCTCCCTCGTGTACCTTATCGGACACAAG 900  
 DB 841 GGCACAAAGTGTAGATGCGCTGAAATCAGGCGCTCCCTCGTGTACCTTATCGGACACAAG 900  
 QY 901 CTCACCTCGGCTATTTGAGGAGCTCTCTGGTGGGTCTAGCCACAGTATCTTTATGCGAAC 960  
 DB 901 CTCACCTCGGCTATTTGAGGAGCTCTCTGGTGGGTCTAGCCACAGTATCTTTATGCGAAC 960  
 QY 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCCGCACTCAACTTGTGCGCAGG 1020  
 DB 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCCGCACTCAACTTGTGCGCAGG 1020  
 QY 1021 TCCAAAGGAGTGATCAATCG 1040  
 DB 1021 TCCAAAGGAGTGATCAATG 1040

RESULT 9  
 ADQ60231  
 ID ADQ60231 standard; DNA; 1041 BP.  
 XX  
 AC ADQ60231;





Fri Nov 12 12:26:24 2004

XX	18-DEC-2003	(first entry)	Db	198	GTGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACGAGGACT	257
DT						
XX		DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 1).	Qy	184	CTAAATACACAGTTTGTATGCTTCTATGGGAGAGNGTACTCAGCAGGACATCTATGCA	243
DE						
XX		human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;	Db	258	CTCAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
KW		cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;	Qy	244	GGTTCAGTGAGCCATCCCTAAAGCAGCTCTGGAAGGCGAGATGCGAGTGTCTTGGCC	303
KW		cancer; hyperplasia; restenosis; cellular proliferation disorder;	Db	318	GGTTCAGTGAGCCATCCCTAAAGCAGCTCTGGAAGGCGAGATGCGAGTGTCTTGGCC	377
XX		cardiac hypertrophy; immune disorder; inflammation.				
OS		Homo sapiens.	Qy	304	TATGGAACCCACAGAGAGCTGGGAAGACGACACAATGTGGGAGAGCCAGAGCAACCTGGG	363
XX			Db	378	TATGGAACCCACAGAGAGCTGGGAAGACGACACAATGTGGGAGAGCCAGAGCAACCTGGG	437
FH		Location/Qualifiers	Qy	364	GTGATCCCGCGGGCTCTCATGGACCTCTGAGCAGCTCACAGGAGAGGGTGGCGAGGGC	423
FT		1. .1115	Db	438	GTGATCCCGCGGGCTCTCATGGACCTCTGAGCAGCTCACAGGAGAGGGTGGCGAGGGC	497
FT		/*tag= a	Qy	424	CGGCCATGGGCCCTTTCTGTCAACCATGTCTTACCTAGAGATCTACAGGAGAGGATTTA	483
FT		/product= "KID protein"	Db	498	CGGCCATGGGCCCTTTCTGTCAACCATGTCTTACCTAGAGATCTACAGGAGAGGATTTA	557
FT		/transl_except= (pos: 1. .5; aa: Met)	Qy	484	GACCTCTCTGGACCTCTTCTGGGAGACCTGTGTAATCCGAGAAGACTGCGGGGGAAATATC	543
FT		/note= "This codon has an apparent 2 nucleotide insertion	Db	558	GACCTCTCTGGACCTCTTCTGGGAGACCTGTGTAATCCGAGAAGACTGCGGGGGAAATATC	617
FT		that alters the reading frame"	Qy	544	CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTGTGATTTTGGAGCGGCACTTC	603
XX			Db	618	CTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTGTGATTTTGGAGCGGCACTTC	677
XX		US6387644-B1.	Qy	604	CTGCCAGGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGTCTCTCCGCG	663
XX		14-MAY-2002.	Db	678	CTGCCAGGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGTCTCTCCGCG	737
XX		28-NOV-2000; 2000US-00724224.	Qy	664	AGTCATGTGTCTCTCGTCAAGGTGACACAGCGGGAACGTTTGGCCCCCATTTGCGCCAG	723
XX		WPI; 2003-706919/67.	Db	738	AGTCATGTGTCTCTCGTCAAGGTGACACAGCGGGAACGTTTGGCCCCCATTTGCGCCAG	797
XX		P-PSDB; ADC23338.	Qy	724	CGAGAGGGAACAACTCTACCTGATTTGACTTGGCTCAGAGGACAAACCGCGCACAGGC	783
XX		Identifying a candidate agent as modulator of function of a target	Db	798	CGAGAGGGAACAACTCTACCTGATTTGACTTGGCTCAGAGGACAAACCGCGCACAGGC	857
XX		protein for treating cellular proliferation disorders by adding a	Qy	784	AAACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTCTGGGC	843
XX		candidate agent to a mixture of the target protein that	Db	858	AAACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTCTGGGC	917
XX		directly/indirectly produces ADP or phosphate.	Qy	844	AAAGTGTAGATGCGCTGAATCAGGGCTCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC	903
XX		Disclosure; SEQ ID NO 1; 26pp; English.	Db	918	AAAGTGTAGATGCGCTGAATCAGGGCTCTCCCTCGTGTACCTTATCGGGACAGCAAGCTC	977
XX			Qy	904	ACTCGCTATTTCAGAGCTCTCTCGGTGGCTCAGCCCAAGTATCTTATTCGCAACATTT	963
XX		This invention relates to a novel method for high throughput screening	Db	978	ACTCGCTATTTCAGAGCTCTCTCGGTGGCTCAGCCCAAGTATCTTATTCGCAACATTT	1037
XX		systems used to identify compounds for the treatment of cellular	Qy	964	GGCCCTTGAGAGAGCGTTTCTACCTAGACAGCTCTCGGACACTCAACTTGTCTGCGAGTCC	1023
XX		proliferation disorders. Specifically, it refers to candidate agents that	Db	1038	GGCCCTTGAGAGAGCGTTTCTACCTAGACAGCTCTCGGACACTCAACTTGTCTGCGAGTCC	1097
XX		are capable of modulating the activity of target proteins having motor	Qy	1024	AGGAGGTGATCAATCG	1040
XX		domains, such that the target protein directly or indirectly produces ADP	Db	1098	AAGGAGGTGATCAATCG	1114
XX		or phosphate. Furthermore, this activity can be determined using				
XX		fluorescence or absorbance readouts. The present invention describes a				
XX		method that identifies modulators of the target protein, which is a				
XX		kinesin-like DNA binding protein (known as KID) as cytostatic, cardiants,				
XX		immunomodulators and antiinflammatories. Accordingly, through gene				
XX		therapy, they can be used for the treatment of cancer, hyperplasias,				
XX		restenosis, cardiac hypertrophy, immune disorders and inflammation. This				
XX		polynucleotide sequence is human KID DNA (SeqID 1) encoding a full length				
XX		KID enzyme of the invention.				
XX						
XX		Query Match				
XX		Best Local Similarity				
XX		Matches 1036; Conservative				
XX		70.7%; Score 1035.4; DB 10; Length 1115;				
XX		99.9%; Pred. No. 2.2e-277;				
XX		0; Mismatches 1; Indels 0; Gaps 0;				
XX						
XX		4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG				
XX		63				
XX		78 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG				
XX		137				
XX		64 GTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG				
XX		123				
XX		138 GTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACCTCCAGCTCGGTAAGG				
XX		197				
XX		124 GTGGGGGCGATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACGAGGACT				
XX		183				

RESULT 11  
ADQ60227  
ID ADQ60227 standard; DNA; 1115 BP.  
XX  
AC ADQ60227;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Human microtubule motor protein DNA #1.



XX Human; microtubule motor protein; gene; ds;  
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
 KW cardiac hypertrophy; immune disorder; inflammation;  
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;  
 KW graft rejection; inflammatory bowel disease; angioplasty.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US672043-B1.  
 XX 13-JUL-2004.  
 XX  
 XX 06-MAR-2002; 2002US-00093317.  
 XX 20-APR-1999; 99US-00295612.  
 XX 20-JUN-2000; 2000US-00597292.  
 XX 28-NOV-2000; 2000US-00724224.  
 XX  
 XX (CYTO-) CYTOKINETICS INC.  
 XX  
 XX Beraud C;  
 XX  
 XX WPI; 2004-532491/51.  
 XX  
 XX New isolated microtubule motor protein, useful for screening modulators  
 PT for treating cellular proliferation disorders such as cancer,  
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.  
 XX  
 XX Disclosure; SEQ ID NO 1; 26pp; English.  
 XX  
 CC The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase  
 CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents DNA encoding a human microtubule  
 CC motor protein of the invention. Note: The specification states that this  
 CC sequence encodes the protein featured as SEQ ID NO:2, but this does not  
 CC appear to be the case.  
 XX  
 SQ Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;  
 Query Match 70.7%; Score 1035.4; DB 12; Length 1115;  
 Best Local Similarity 99.9%; Pred. No. 2.2e-277;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GGTGCTGTGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCGAGTCCGCTAAGG 63  
 DB 78 GGTGCTGTGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCGAGTCCGCTAAGG 137  
 QY 64 GTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCGAGTCCGCTAAGG 123  
 DB 138 GTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCGAGTCCGCTAAGG 197  
 QY 124 GTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCGAGTCCGCTAAGG 183  
 DB 198 GTGGCTGTGGCTAAGCAAGATTGGAGTACTCGTCTGCTCACTCGAGTCCGCTAAGG 257  
 QY 184 CTCAATACCAATTTGATGCTTCTATGTTGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243

Db 258 CTCAATACCAATTTGATGCTTCTATGTTGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317  
 QY 244 GGTTCAGTGCAGCCCATCTTAAGCAGTCTGCTGGAAGGCGAGAAATGCCAGTGTGCTGCC 303  
 DB 318 GGTTCAGTGCAGCCCATCTTAAGCAGTCTGCTGGAAGGCGAGAAATGCCAGTGTGCTGCC 377  
 QY 304 TATGGACCCCAAGAGAGTGGGAAGACGACACAATGCTGGGCGAGCCCAAGAGCAACTGGG 363  
 DB 378 TATGGACCCCAAGAGAGTGGGAAGACGACACAATGCTGGGCGAGCCCAAGAGCAACTGGG 437  
 QY 364 GTGATCCCGGGGCTCTCATGGAACCTCTGTCAGTCTCAAGAGGAGGAGGTGCCAGGGC 423  
 DB 438 GTGATCCCGGGGCTCTCATGGAACCTCTGTCAGTCTCAAGAGGAGGAGGTGCCAGGGC 497  
 QY 424 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGTCTACCAAGGAGAAAGTATTA 483  
 DB 498 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGTCTACCAAGGAGAAAGTATTA 557  
 QY 484 GACCTCTGGACCCCTGCTTCCGGGAGAGCTGTGTAATCCGAGAAAGTGTCCGGGGAATATC 543  
 DB 558 GACCTCTGGACCCCTGCTTCCGGGAGAGCTGTGTAATCCGAGAAAGTGTCCGGGGAATATC 617  
 QY 544 CTGATTCGGGCTCTCTCCAGAACCCCATCAGTAGTCTTGTGATTTTGAAGCGGCACTTC 603  
 DB 618 CTGATTCGGGCTCTCTCCAGAACCCCATCAGTAGTCTTGTGATTTTGAAGCGGCACTTC 677  
 QY 604 CTGCCAGCCAGTCCGAATCCGAGCTGTAGGAGCCACCCGGCTCAACCAAGCGCTCTCCCGC 663  
 DB 678 CTGCCAGCCAGTCCGAATCCGAGCTGTAGGAGCCACCCGGCTCAACCAAGCGCTCTCCCGC 737  
 QY 664 AGTCATCTGTGCTCTCTGTCGAGTGGACAGCGGGAACGTTTGGCCCCCAATTTCCGCCAG 723  
 DB 738 AGTCATCTGTGCTCTCTGTCGAGTGGACAGCGGGAACGTTTGGCCCCCAATTTCCGCCAG 797  
 QY 724 CGAGAGGAAAACCTCTACCTGATTTGACTTGGTGGGTGAGAGGACAAACCGGCGCACAGGC 783  
 DB 798 CGAGAGGAAAACCTCTACCTGATTTGACTTGGTGGGTGAGAGGACAAACCGGCGCACAGGC 857  
 QY 784 AACAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCTGGGC 843  
 DB 858 AACAGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCTGGGC 917  
 QY 844 AAAGTGTAGATGCGCTGAATCAAGGGCTCTCTGCTGCTACCTTATCGGAGCAGCAAGCTC 903  
 DB 918 AAAGTGTAGATGCGCTGAATCAAGGGCTCTCTGCTGCTACCTTATCGGAGCAGCAAGCTC 977  
 QY 904 ACTGCGCTATTGAGAGTCTCTGCGGTGGTCAAGCCCAAGATATCTTATTTGCCAACATT 963  
 DB 978 ACTGCGCTATTGAGAGTCTCTGCGGTGGTCAAGCCCAAGATATCTTATTTGCCAACATT 1037  
 QY 964 GCCCTGAGAGCGCTTCTACCTAGACAGTCTCCGACCTCAACTTTGCTGCCAGTCC 1023  
 DB 1038 GCCCTGAGAGCGCTTCTACCTAGACAGTCTCCGACCTCAACTTTGCTGCCAGTCC 1097  
 QY 1024 AAGGAGGTGATCAATCG 1040  
 DB 1098 AAGGAGGTGATCAATCG 1114

RESULT 12  
 ABK70260  
 ID ABK70260 standard; cDNA; 491 BP.  
 XX  
 AC ABK70260;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human lung cancer associated cDNA SEQ ID 131.  
 XX  
 KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.  
 XX  
 OS Homo sapiens.

XX WO200224057-A2.  
 PN 28-MAR-2002.  
 XX 20-SEP-2001; 2001WO-US042232.  
 XX 22-SEP-2000; 2000US-0234837P.  
 PR 10-OCT-2000; 2000US-0239440P.  
 PR 29-JUN-2001; 2001US-0301928P.  
 XX (CORI-) CORIXA CORP.  
 XX Benson DR, Mohamath R, Lodes MJ;  
 XX WPI; 2002-372001/40:  
 XX New tumor lung proteins and nucleic acids encoding the proteins, useful  
 PT as vaccines and for treating, preventing, diagnosing or monitoring lung  
 PT cancer.  
 XX  
 PS Claim 1; Page 144; 189pp; English.  
 XX The invention relates to an isolated polynucleotide comprising a sequence  
 CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),  
 CC or their fragments, homologues, variants or complements and their encoded  
 CC polypeptides. Also included are an expression vector comprising the  
 CC polynucleotide operably linked to an expression control sequence; a host  
 CC cell transformed or transfected with an expression vector of; an isolated  
 CC antibody, or its antigen-binding fragment that specifically binds to the  
 CC polypeptide; a method for detecting the presence of a cancer in a patient  
 CC ; a fusion protein comprising at least the polypeptide; an  
 CC oligonucleotide that hybridises to the polynucleotide under moderately  
 CC stringent conditions; a method for stimulating and/or expanding T cells  
 CC specific for a tumor protein; an isolated T cell population comprising a first  
 CC cells prepared from the method of above; a composition comprising a first  
 CC component consisting of carriers and immunostimulants, and a second  
 CC component selected from the polynucleotides, proteins, antibodies, fusion  
 CC proteins, T cell populations and antigen presenting cells expressing the  
 CC polypeptide; methods for stimulating an immune response or treating  
 CC cancer in a patient by administering the composition and diagnostic kits  
 CC comprising at least one of the oligonucleotide of, or an antibody and a  
 CC detection reagent consisting of a reporter group. The polypeptides and  
 CC polynucleotides are useful as vaccines for the treatment or prevention of  
 CC lung cancer, and for diagnosis and monitoring of such cancer. The  
 CC polynucleotide, polypeptide and antigen presenting cells can be used to  
 CC stimulate or expand T cells specific for a tumorous protein. The  
 CC polynucleotides may be used as probes or primers for nucleic acid  
 CC hybridisation, and in the preparation of ribozyme molecules for  
 CC inhibiting expression of tumour polypeptides and proteins in tumour  
 CC cells. The present sequence is one of the 183 lung cancer associated  
 CC polynucleotides  
 XX  
 SQ Sequence 491 BP; 127 A; 140 C; 136 G; 88 T; 0 U; 0 Other;

Db 180 ACCAGAGGCAAGAGAGCGCCCTGAGAGGAGGAGATTGGAGCCCTGAGCCCAT 239  
 Qy 1176 GGCAGCTCCAGCTCTGCTCCAGAGAACTCAGCCCCCTACAGAGCTAAGCAGCATGGA 1235  
 Db 240 GGCAGCTCCAGCTCTGCTCCAGAGAACTCAGCCCCCTACAGAGCTAAGCAGCATGGA 299  
 Qy 1236 CCGGGCCATGCTGGAGCGCTCTCTCAGCTTGGAGCCGCTCTGCTTGCCTCCAGGGAGCCCA 1295  
 Db 300 CCGGGCCATGCTGGAGCGCTCTCTCAGCTTGGAGCCGCTCTGCTTGCCTCCAGGGAGCCCA 359  
 Qy 1296 GGGGGCCCTCTGTTGAGTATACCCCAAGCGAGAGCGGATGGTGTATAGACAGTAGA 1355  
 Db 360 GGGGGCCCTCTGTTGAGTATACCCCAAGCGAGAGCGGATGGTGTATAGACAGTAGA 419  
 Qy 1356 AGAGAGGACCTAGAGATTGAGAGCTTAAGACGAGCAAGCAAAAGAACTGGAGGCCAAGAT 1415  
 Db 420 AGAGAGGACCTAGAGATTGAGAGCTTAAGACGAGCAAGCAAAAGAACTGGAGGCCAAGAT 479  
 Qy 1416 GTTGGCCCGAGAA 1427  
 Db 480 GTTGGCCCGAGAA 491  
 RESULT 13  
 AAS38879  
 ID AAS38879 standard; cDNA; 386 BP.  
 XX AAS38879;  
 AC AAS38879;  
 XX 17-DEC-2001 (first entry)  
 DT Novel human diagnostic and therapeutic gene #1937.  
 XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.  
 KW Homo sapiens.  
 OS WO200166753-A2.  
 PN 13-SEP-2001.  
 PD 09-MAR-2001; 2001WO-US0007787.  
 PF 09-MAR-2000; 2000US-0188609P.  
 PR (CHIR ) CHIRON CORP.  
 PA (HYSE-) HYSEQ INC.  
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;  
 PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
 PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Grain B;  
 XX WPI; 2001-530177/58.  
 XX New polynucleotides and polypeptides, useful for diagnosis and treatment  
 PT of breast, lung and colon cancer.  
 XX  
 XX Claim 1; Page 1080; 1193pp; English.  
 XX The invention relates to new polynucleotides and polypeptides, useful for  
 CC diagnosis and treatment of breast, lung and colon cancer. The sequences  
 CC can be used in detecting differentially expressed genes correlated with a  
 CC cancerous state of a mammalian cell, comprising detecting at least one  
 CC differentially expressed gene product in a test sample derived from a  
 CC cell suspected of being cancerous. They can also be used to inhibit  
 CC tumour growth by modulating expression of a gene product. AAS36943-  
 CC AAS39338 represent novel human diagnostic and therapeutic coding  
 CC sequences of the invention  
 XX  
 SQ Sequence 386 BP; 105 A; 102 C; 118 G; 61 T; 0 U; 0 Other;  
 Query Match 25.7%; Score 376.4; DB 4; Length 386;

Best Local Similarity 98.4%; Pred. No. 3.2e-94;  
Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1073 CTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACGAGAGCAAGAGAG 1132  
Db 1 CGTTGCTGCTGTTAAGCTGTCTCAGAAAGAAATTCCTTGGTCCACGAGAGCAAGAGAG 60

QY 1133 CCCGAGGCTTGAAGAGAGAGATTTGGAGCCCTTGAGCCCATGCGAGCTCCAGCCTCTG 1192  
Db 61 CCCGAGGCTTGAAGAGAGAGATTTGGAGCCCTTGAGCCCATGCGAGCTCCAGCCTCTG 120

QY 1193 CTTCCCAAACTAGCCCTTACAGAGCTTAAGCAGATGAGCCGCGCATGCTGGAGC 1252  
Db 121 CTTCCCAAACTAGCCCTTACAGAGCTTAAGCAGATGAGCCGCGCATGCTGGAGC 180

QY 1253 GCCTCTCAGCTTGGACCGCTGCTGCTCCAGGGGAGCCAGGGGCGCCCTCTGTTGA 1312  
Db 181 GCCTCTCAGCTTGGACCGCTGCTGCTCCAGGGGAGCCAGGGGCGCCCTCTGTTGA 240

QY 1313 GTACCCCAAGCAGAGCGGATGTGCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGA 1372  
Db 241 GTACCCCAAGCAGAGCGGATGTGCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGA 300

QY 1373 TTGAGAGCTTAAAGACGAGCAAAAGAACTGGAGGCCCAAGATGTTGGCCCAAGAGGCTG 1432  
Db 301 TTGAGAGCTTAAAGACGAGCAAAAGAACTGGAGGCCCAAGATGTTGGCCCAAGAGGCTG 360

QY 1433 AGGAAAGGAGAACCATTTGCTCCCA 1458  
Db 361 AGGAAAGGAGAACCATTTGCTCCCA 386

RESULT 14  
ACH44493  
ID ACH44493 standard; cDNA; 464 BP.  
XX ACH44493;  
XX 13-OCT-2003 (first entry)  
XX Human foetal brain cDNA #5218.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX 17-APR-2003.  
XX 30-JUL-2001; 2001US-00918995.  
XX 30-JUL-2001; 2001US-00918995.  
XX (DRMA//) DRMANAC R T.  
XX (LABA//) LABAT I.  
XX (STAC//) STACHE-CRAIN B.  
XX (DICK//) DICKSON M C.  
XX (JONE//) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
XX as hybridization probes, as oligomers for PCR, for chromosome and gene  
XX mapping, in the recombinant production of protein, or in generating  
XX antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 31705; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of

CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030073623  
XX  
SQ Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;

Query Match 18.1%; Score 265; DB 9; Length 464;  
Best Local Similarity 77.6%; Pred. No. 3.5e-63;  
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 126 GCGGGGATGGACAGCTGCTCTTAGAGATTGCTAAGTGGAGGAAACCAACGAGGACTCT 185  
Db 37 GCCTGGAATTCATATATCGGTCGATAGAGTGTAAACGGCCGCAACCAACCGATAATCT 96

QY 186 CAATATACAGTTTATGCTTCTATGGGAGAGAGTACTCAGCAGAGCATCTATGCAAG 245  
Db 97 CCCTTACCTTGCTTGCCTTTGACAGCCATAGGAGGACTGCGCTTGACATGTACCTTCC 156

QY 246 TTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTGCCTA 305  
Db 157 TTGACCGTGGACCGTGCACCTGCCACTGCTCTTCGTAAATAAGGAGGAGCTGGCCTA 216

QY 306 TGGACCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCCGAGCAACCTGGGGT 365  
Db 217 CGGTGCTGAGGAGCTGGCATTAAC-ATCTGATGCTTGGCAGCCCGAGCAACCTGGGGT 275

QY 366 GATCCCGGGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGCGC 425  
Db 276 GATCCCGGGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGCGC 335

QY 426 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCCAGAGAGGTATTAGA 485  
Db 336 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCCAGAGAGGTATTAGA 395

QY 486 CCTCCTGGACCTGCTTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAATATCCT 545  
Db 396 CCTCCTGGACCTGCTTTCGGGAGACCTTGGTAATCCGAGAGACTGCGGGGGAATATCCT 455

QY 546 GATTCCGGG 554  
Db 456 GATTCCGGG 464

RESULT 15  
ACH77408  
ID ACH77408 standard; DNA; 531 BP.  
XX ACH77408;  
XX 29-JUL-2004 (first entry)  
XX  
XX Human genome derived single exon probe #10603.  
XX Human; probe; ss; gene expression; single exon probe; microarray;  
XX alternative splicing event; genomic alteration.  
XX  
XX Homo sapiens.  
XX  
XX US2003194704-A1.

Db 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGGCAAGTGGTAGATGCGCTGAA 213  
QY 864 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTGCGAGAC 921  
Db 214 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTCGCCTATTGCGAGTC 271

Search completed: November 10, 2004, 06:58:00  
Job time : 760.977 secs

PD 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX  
XX New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
XX Claim 15; SEQ ID NO 10603; 80pp; English.  
XX  
CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressable or isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX  
XX Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;  
SQ  
Query Match 15.9%; Score 233.2; DB 12; Length 531;  
Best Local Similarity 98.7%; Pred. No. 2.6e-54;  
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 684 CAAGGTGGACACGGGAACTTTGGCCCATTTCCGACGAGAGGAAACTCTACCT 743  
Db 34 CCAGGTGGACACGGGAACTTTGGCCCATTTCCGACGAGAGGAAACTCTACCT 93  
QY 744 GATTGACTTGGCTGGGTACAGGACACCGCGCACAGGCAACAAGGGCTTCGGCTAAA 803  
Db 94 GATTGACTTGGCTGGGTACAGGACACCGCGCACAGGCAACAAGAGCCTTCGGCTAAA 153  
QY 804 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCTGGGCAAGTGGTAGATGCGCTGAA 863

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 75.9417 Seconds  
(without alignments)  
1747.786 Million cell updates/sec

Title: US-10-797-893-2

Perfect score: 1878

Sequence: 1 MAAGSTQQRREMAAASAA.....FYLDTVSALNFAARSKEVIN 370

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 359729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	370	7	Adc23338 Human kin
2	1878	100.0	370	8	Adg60228 Human mic
3	1878	100.0	490	7	Adk40973 Novel hum
4	1878	100.0	665	8	Adg09240 Human KNS
5	1873	98.7	512	7	Adc23340 Human kin
6	1873	99.7	512	8	Adg60230 Human mic
7	1763	93.9	346	7	Adc23342 Human kin
8	1763	93.9	346	8	Adg60232 Human mic
9	1763	93.9	460	3	Aab56650 Human pro
10	1763	93.9	487	7	Adc23344 Human kin
11	1763	93.9	487	8	Adg60234 Human mic
12	559.5	29.8	784	4	Abb71112 Drosophil
13	559	29.8	548	7	Adm04007 Human pro
14	559	29.8	998	8	Ado44167 Structura
15	552.5	29.4	357	5	Aau76958 Novel hum
16	552.5	29.4	898	4	Abu53123 Intracell
17	552.5	29.4	898	5	Aau76957 Novel hum
18	552.5	29.4	898	7	Adc31082 Human nov
19	552.5	29.4	898	7	Adi15915 Human pp
20	551.5	29.4	603	6	Abu11606 Human MDD
21	549	29.2	905	5	Aau76967 Novel hum
22	541	28.8	408	8	Adn40550 Candida a
23	541	28.8	408	8	Adn40554 Candida a
24	541	28.8	972	8	Adn40548 Candida a
25	541	28.8	974	8	Adn40552 Candida a

26	535.5	28.5	677	4	Abb65183 Drosophil
27	528	28.1	383	5	Abb07412 Amino aci
28	528	28.1	383	6	Abg72053 Human Hsk
29	528	28.1	864	5	Abb07410 Human kin
30	528	28.1	864	6	Abg72052 Human kin
31	526	28.0	375	5	Aae14401 Human Hsk
32	526	28.0	409	5	Aae14402 Human Hsk
33	526	28.0	409	5	Aau79592 Human Hsk
34	526	28.0	419	7	Adc31540 Human nov
35	526	28.0	1388	5	Aae14400 Human kin
36	526	28.0	1388	5	Aau79590 Human kin
37	526	28.0	1388	6	AbR48222 Human bla
38	526	28.0	1388	7	AdB80468 Ovarian c
39	526	28.0	1388	7	Adc35116 Human bre
40	526	28.0	1388	8	AdL83290 Human pro
41	526	28.0	1388	8	Adq20128 Human sof
42	526	28.0	1388	8	Adq09226 Human KNS
43	525.5	28.0	757	4	Aau19569 Human dia
44	525.5	28.0	757	5	Abp51294 Human MDD
45	523.5	27.9	834	7	Adj95078 Novel NOV

## ALIGNMENTS

RESULT 1  
ID ADC23338 standard; protein; 370 AA.  
XX AC ADC23338;  
XX DT 18-DEC-2003 (first entry)  
XX DE Human kinesin-like DNA binding protein (KID) (SeqID 2).  
XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
XX KW cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy;  
XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
XX KW cardiac hypertrophy; immune disorder; inflammation.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 1 /note= "Encoded by ATGCA"  
XX PN US6387644-B1.  
XX PD 14-MAY-2002.  
XX PF 28-NOV-2000; 2000US-00724224.  
XX PR 20-APR-1999; 99US-00295612.  
XX PR 20-JUN-2000; 2000US-00597292.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C;  
XX DR WPI; 2003-706919/67.  
XX DR N-PSDB; ADC23337.  
XX PT Identifying a candidate agent as modulator of function of a target  
XX PT protein for treating cellular proliferation disorders by adding a  
XX PT candidate agent to a mixture of the target protein that  
XX PT directly/indirectly produces ADP or phosphate.  
XX PS Claim 1; SEQ ID NO 2; 26pp; English.  
XX CC This invention relates to a novel method for high throughput screening  
XX CC systems used to identify compounds for the treatment of cellular  
XX CC proliferation disorders. Specifically, it refers to candidate agents that  
XX CC are capable of modulating the activity of target proteins having motor

CC domains, such that the target protein directly or indirectly produces ADP  
CC or phosphate. Furthermore, this activity can be determined using  
CC fluorescence or absorbance readouts. The present invention describes a  
CC method that identifies modulators of the target protein, which is a  
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardiantes,  
CC immunomodulators and antiinflammatory. Accordingly, through gene  
CC therapy, they can be used for the treatment of cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
CC polypeptide sequence is human KID protein (SeqID 2) of the invention.  
XX  
XX  
SQ Sequence 370 AA;

Query Match 100.0%; Score 1878; DB 7; Length 370;  
Best Local Similarity 100.0%; Pred. No. 3.8e-182;  
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAAGGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPPVDGTAGA 60  
Db 1 MAAGGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPPVDGTAGA 60  
Qy 61 SDPPCVRGMDSCSLEIANRNHQTLYQFDIFYGERSTQDDIYAGSVQPIRLHLLGQN 120  
Db 61 SDPPCVRGMDSCSLEIANRNHQTLYQFDIFYGERSTQDDIYAGSVQPIRLHLLGQN 120  
Qy 121 ASVLAYGPTGAGKTHMLGSPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180  
Db 121 ASVLAYGPTGAGKTHMLGSPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180  
Qy 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRRTVGATRLN 240  
Db 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRRTVGATRLN 240  
Qy 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSDNRRTGNKGLKESGAINTS 300  
Db 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSDNRRTGNKGLKESGAINTS 300  
Qy 301 LFLVGKVVDAALNOGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERFYLDTVSALN 360  
Db 301 LFLVGKVVDAALNOGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERFYLDTVSALN 360  
Qy 361 FAARKEVIN 370  
Db 361 FAARKEVIN 370

RESULT 2  
ADQ60228  
ID ADQ60228 standard; protein; 370 AA.  
AC ADQ60228;  
XX  
XX  
XX 23-SEP-2004 (first entry)  
XX Human microtubule motor protein #1.  
XX Human; microtubule motor protein; cellular proliferation disorder;  
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
XX  
XX Homo sapiens.  
XX US6762043-B1.  
XX 13-JUL-2004.  
XX  
XX 06-MAR-2002; 2002US-00093317.  
XX  
XX 20-APR-1999; 99US-00295612.  
XX 20-JUN-2000; 2000US-00597252.  
XX 28-NOV-2000; 2000US-0072424.  
XX  
XX (CYTO-) CYTOKINETICS INC.

XX  
PI Beraud C;  
XX  
XX WPI; 2004-532491/51.  
XX  
XX New isolated microtubule motor protein, useful for screening modulators  
PT for treating cellular proliferation disorders such as cancer,  
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
PT inflammation.  
XX  
XX  
PS Claim 1; SEQ ID NO 2; 26pp; English.

XX The invention relates to human microtubule motor proteins and the nucleic  
CC acids encoding them. The invention also relates to a method of screening  
CC for modulators of a motor protein which has microtubule stimulated ATPase  
CC activity, a method of testing for ATPase activity of microtubule motor  
CC proteins, methods to identify candidate agents that bind to a target  
CC protein or act as a modulator of the binding characteristics or  
CC biological activity of a target protein, modulators of the target  
CC protein, and methods of treating cellular proliferation disorders such as  
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
CC and inflammation, for treating disorders associated with kinesin-like DNA  
CC binding protein (KID) and for inhibiting KID. The sequences are used for  
CC screening for modulators of motor proteins useful for treating cellular  
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
CC hypertrophy, immune disorders and inflammation, for treating disorders  
CC associated with KID and for inhibiting KID and for treating autoimmune  
CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents a human microtubule motor protein  
CC of the invention. Note: The specification states that this sequence is  
CC encoded by the nucleic acid featured as SEQ ID NO:1, but this does not  
CC appear to be the case.

XX SQ Sequence 370 AA;  
Query Match 100.0%; Score 1878; DB 8; Length 370;  
Best Local Similarity 100.0%; Pred. No. 3.8e-182;  
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAGGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPPVDGTAGA 60  
Db 1 MAAGGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPARVRVAVRLRPPVDGTAGA 60  
Qy 61 SDPPCVRGMDSCSLEIANRNHQTLYQFDIFYGERSTQDDIYAGSVQPIRLHLLGQN 120  
Db 61 SDPPCVRGMDSCSLEIANRNHQTLYQFDIFYGERSTQDDIYAGSVQPIRLHLLGQN 120  
Qy 121 ASVLAYGPTGAGKTHMLGSPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180  
Db 121 ASVLAYGPTGAGKTHMLGSPQPGVPRALMDLLQLTREGAEGRPWALSVTMSYLEIY 180  
Qy 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRRTVGATRLN 240  
Db 181 QEKVLDLLDPASGDLVIREDCRGNILIPGLSQPISSFADFPERHFLPASNRRTVGATRLN 240  
Qy 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSDNRRTGNKGLKESGAINTS 300  
Db 241 QRSRSHAVLLVKVDQERLAPFRQREGKLYLIDLAGSDNRRTGNKGLKESGAINTS 300  
Qy 301 LFLVGKVVDAALNOGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERFYLDTVSALN 360  
Db 301 LFLVGKVVDAALNOGLPRVPYRDSKLTLLQDSLGGSAHSILIANIAPERFYLDTVSALN 360  
Qy 361 FAARKEVIN 370  
Db 361 FAARKEVIN 370

RESULT 3  
ADK40973  
ID ADK40973 standard; protein; 490 AA.  
XX



CC the expression of the THAP responsive gene. Also described: (1) a method  
CC of modulating the expression of a gene responsive to a THAP/chemokine  
CC complex; (2) a pharmaceutical composition comprising a THAP responsive  
CC element in a pharmaceutical carrier; (3) a transcription factor decoy  
CC consisting essentially of a THAP responsive element; (4) a cell  
CC comprising a transcription factor decoy described above; (5) methods of  
CC modulating the interaction between a nucleic acid and a THAP-family  
CC polypeptide or its biological fragment, or a nucleic acid and a  
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a  
CC cell comprising a viral vector which comprises a promoter operably linked  
CC to a nucleic acid encoding a THAP-family polypeptide or its biological  
CC fragment; (7) a method of constructing a cell which expresses a  
CC recombinant THAP-family polypeptide; (8) a method of ameliorating  
CC symptoms associated with a condition mediated by a THAP/chemokine complex  
CC ; (9) methods of identifying a test compound that modulates transcription  
CC at a THAP responsive element or that modulates the transport of a  
CC chemokine into the nucleus; (10) methods for reducing the symptoms  
CC associated with a condition selected from excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative  
CC diseases; symptoms associated with a condition resulting from the  
CC activity of a chemokine or a THAP-family polypeptide in an individual; or  
CC symptoms associated with transcriptional repression or activation  
CC mediated by a THAP-family polypeptide in an individual; (11) a vector  
CC comprising a THAP responsive promoter operably linked to a nucleic acid  
CC encoding a detectable product; (12) a genetically engineered cell  
CC comprising the vector described above or that expresses a THAP-family  
CC polypeptide or its biological fragment; (13) an in vitro transcription  
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,  
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-  
CC family polypeptide that does not bind to a chemokine. The pharmaceutical  
CC composition has antiangiogenic, antiinflammatory, cardiovascular,  
CC cytoskeletal, neuroprotective and osteopathic activities, and can be used  
CC as a THAP and THAP synthesis modulator. The composition can be used for  
CC modulating the expression of a THAP responsive gene. Modulation is useful  
CC for reducing symptoms of conditions such as excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative  
CC diseases. The present sequence is used in the exemplification of the  
CC present invention.

XX Sequence 665 AA;  
Query Match 100.0%; Score 1878; DB 8; Length 665;  
Best Local Similarity 100.0%; Pred. No. 9,5e-182; Indels 0; Gaps 0;  
Matches 370; Conservative 0; Mismatches 0;  
Qy 1 MAAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 60  
Db 1 MAAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 60  
Qy 61 SDPPCVRGMDSCSLEIANWRNHQETLYQFDAPFYGERSTOODIYAGSVQPIRLHLEGON 120  
Db 61 SDPPCVRGMDSCSLEIANWRNHQETLYQFDAPFYGERSTOODIYAGSVQPIRLHLEGON 120  
Qy 121 ASVLAYGTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRGPRWALSVTMSYLEIY 180  
Db 121 ASVLAYGTGAGKTHMLGSPQGVIPRALMDLLQLTREBAGRGPRWALSVTMSYLEIY 180  
Qy 181 QEKVLDDLDPASGDIVIREDCRGNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLN 240  
Db 181 QEKVLDDLDPASGDIVIREDCRGNILIPGLSQKPISSFADFERHFLPASNRRTVGATRLN 240  
Qy 241 QRSRSHAVLLVKVDQRELLAPFRQREKGLYLIDLAGSDNRRTGNKGLRUKESGANTIS 300  
Db 241 QRSRSHAVLLVKVDQRELLAPFRQREKGLYLIDLAGSDNRRTGNKGLRUKESGANTIS 300  
Qy 301 LFLVLGVVDALNOGLPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERFYLDTVSALN 360  
Db 301 LFLVLGVVDALNOGLPRVPYRDSKLTLLQDSLGGSAAHSILIANIAPERFYLDTVSALN 360  
Qy 361 FAARSKEVIN 370  
| | | | |

Db 361 FAARSKEVIN 370  
RESULT 5  
ID ADC23340 standard; protein; 512 AA.  
XX ADC23340;  
AC ADC23340;  
DT 18-DEC-2003 (first entry)  
XX Human kinesin-like DNA binding protein (KID) (SeqID 4).  
DE human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
XX cytoskeletal; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Misc-difference 2 /note= "Encoded by CA"  
FT US6387644-B1.  
XX PN 14-MAY-2002.  
XX PD 28-NOV-2000; 2000US-00724224.  
XX PF 20-APR-1999; 99US-00295612.  
XX PR 20-JUN-2000; 2000US-00597292.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX Beraud C;  
XX WPI: 2003-706919/67.  
XX N-PSDB; ADC23339.  
PT Identifying a candidate agent as modulator of function of a target  
PT protein for treating cellular proliferation disorders by adding a  
PT candidate agent to a mixture of the target protein that  
PT directly/indirectly produces ADP or phosphate.  
XX Claim 1; SEQ ID NO 4; 26pp; English.  
CC This invention relates to a novel method for high throughput screening  
CC systems used to identify compounds for the treatment of cellular  
CC proliferation disorders. Specifically, it refers to candidate agents that  
CC are capable of modulating the activity of target proteins having motor  
CC domains, such that the target protein directly or indirectly produces ADP  
CC or phosphate. Furthermore, this activity can be determined using  
CC fluorescence or absorbance readouts. The present invention describes a  
CC method that identifies modulators of the target protein, which is a  
CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiant,  
CC immunomodulators and antiinflammatory. Accordingly, through gene  
CC therapy, they can be used for the treatment of cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
CC polypeptide sequence is human KID protein (SeqID 4) of the invention.  
XX Sequence 512 AA;  
Query Match 99.7%; Score 1873; DB 7; Length 512;  
Best Local Similarity 100.0%; Pred. No. 2e-181;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 AAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 61  
Db 3 AAGGSTQORREMAAASAAISGAGRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGA 62  
Qy 62 DPFCVRGMDSCSLEIANWRNHQETLYQFDAPFYGERSTOODIYAGSVQPIRLHLEGON 121  
| | | | |



Db 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAYGERSTQDDIYAGSVQPIRLHLLGQNA 122  
 QY 122 SVLAYGPTGAGKTHMTLGSPEQPGVPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 181  
 Db 123 SVLAYGPTGAGKTHMTLGSPEQPGVPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182  
 QY 182 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFERHFLPASNRRTVGATRLNQ 241  
 Db 183 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFERHFLPASNRRTVGATRLNQ 242  
 QY 242 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 301  
 Db 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 302  
 QY 302 FVLGKVVADALNOGLPRVPRYRDSKLTRLQDSLGSSAHSILIANIAPERFYLDTVSALNF 361  
 Db 303 FVLGKVVADALNOGLPRVPRYRDSKLTRLQDSLGSSAHSILIANIAPERFYLDTVSALNF 362  
 QY 362 AARSKEVIN 370  
 Db 363 AARSKEVIN 371  
 RESULT 6  
 ADQ60230  
 ID ADQ60230 standard; protein; 512 AA.  
 XX  
 AC ADQ60230;  
 DT 23-SEP-2004 (first entry)  
 XX  
 DE Human microtubule motor protein #2.  
 KW Human; microtubule motor protein; cellular proliferation disorder;  
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6762043-B1.  
 XX  
 PD 13-JUL-2004.  
 XX  
 XX 06-MAR-2002; 2002US-00093317.  
 XX  
 PR 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 PR 28-NOV-2000; 2000US-00724224.  
 XX  
 PA (CYTO-) CYTOKINETICS INC.  
 XX  
 XX Beraud C;  
 PI  
 DR WPI; 2004-532491/51.  
 XX  
 PT New isolated microtubule motor protein, useful for screening modulators  
 PT for treating cellular proliferation disorders such as cancer,  
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.  
 XX  
 PS Claim 1; SEQ ID NO 4; 26pp; English.  
 XX  
 CC The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase  
 CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA

CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention. Note: The specification states that this sequence is  
 CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not  
 CC appear to be the case.  
 XX  
 SQ Sequence 512 AA;  
 Query Match 99.7%; Score 1873; DB 8; Length 512;  
 Best Local Similarity 100.0%; Pred. No. 2e-181;  
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 AAGSTQQRREMAAASAAAIISGACRCLSKIGATRRPPARVRVAVRLRPPFVDGTAGAS 61  
 Db 3 AAGSTQQRREMAAASAAAIISGACRCLSKIGATRRPPARVRVAVRLRPPFVDGTAGAS 62  
 QY 62 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAYGERSTQDDIYAGSVQPIRLHLLGQNA 121  
 Db 63 DPPCVRGMDSCSLEIANWRNHQETLKYQDFAYGERSTQDDIYAGSVQPIRLHLLGQNA 122  
 QY 122 SVLAYGPTGAGKTHMTLGSPEQPGVPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 181  
 Db 123 SVLAYGPTGAGKTHMTLGSPEQPGVPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQ 182  
 QY 182 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFERHFLPASNRRTVGATRLNQ 241  
 Db 183 EKVLDDLPASGDLVIREDCRGNILPGLSQPISSFADFERHFLPASNRRTVGATRLNQ 242  
 QY 242 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 301  
 Db 243 RSSRSHAVLLVKVDQERLAPFRQEGKLYLIDLAGSEDNRTGNKGLKESGAINTSL 302  
 QY 302 FVLGKVVADALNOGLPRVPRYRDSKLTRLQDSLGSSAHSILIANIAPERFYLDTVSALNF 361  
 Db 303 FVLGKVVADALNOGLPRVPRYRDSKLTRLQDSLGSSAHSILIANIAPERFYLDTVSALNF 362  
 QY 362 AARSKEVIN 370  
 Db 363 AARSKEVIN 371  
 RESULT 7  
 ADC23342  
 ID ADC23342 standard; protein; 346 AA.  
 XX  
 AC ADC23342;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human kinesin-like DNA binding protein (KID) (SeqID 6).  
 XX  
 KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
 KW cytoskeletal; cardiac; immunomodulator; antiinflammatory; gene therapy;  
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
 KW cardiac hypertrophy; immune disorder; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6387644-B1.  
 XX  
 PD 14-MAY-2002.  
 XX  
 PF 28-NOV-2000; 2000US-00724224.  
 XX  
 PR 20-APR-1999; 99US-00295612.  
 PR 20-JUN-2000; 2000US-00597292.  
 XX

PA	(CYTO-) CYTOKINETICS INC.	XX	Homo sapiens.	XX
XX	Beraud C;	XX	US6762043-B1.	XX
PI		XX	13-JUL-2004.	XX
DR	WPI; 2003-706919/67.	XX	06-MAR-2002; 2002US-00093317.	XX
DR	N-PSDB; ADC23341.	XX	20-APR-1999; 99US-00295612.	XX
XX	Identifying a candidate agent as modulator of function of a target	XX	20-JUN-2000; 2000US-00597292.	XX
PT	protein for treating cellular proliferation disorders by adding a	XX	28-NOV-2000; 2000US-00724224.	XX
PT	candidate agent to a mixture of the target protein that	XX	(CYTO-) CYTOKINETICS INC.	XX
PT	directly/indirectly produces ADP or phosphate.	XX	Beraud C;	XX
XX	Claim 1; SEQ ID NO 6; 26pp; English.	XX	WPI; 2004-532491/51.	XX
PS		XX	N-PSDB; ADQ60231.	XX
XX	This invention relates to a novel method for high throughput screening	XX	New isolated microtubule motor protein, useful for screening modulators	XX
CC	systems used to identify compounds for the treatment of cellular	XX	for treating cellular proliferation disorders such as cancer,	XX
CC	proliferation disorders. Specifically, it refers to candidate agents that	XX	hyperplasias, restenosis, cardiac hypertrophy, immune disorders and	XX
CC	are capable of modulating the activity of target proteins having motor	XX	inflammation.	XX
CC	domains, such that the target protein directly or indirectly produces ADP	XX		XX
CC	or phosphate. Furthermore, this activity can be determined using	XX	Claim 1; SEQ ID NO 6; 26pp; English.	XX
CC	fluorescence or absorbance readouts. The present invention describes a	XX	The invention relates to human microtubule motor proteins and the nucleic	XX
CC	method that identifies modulators of the target protein, which is a	XX	acids encoding them. The invention also relates to a method of screening	XX
CC	kinesin-like DNA binding protein (known as KID) as cytostatic, cardants,	XX	for modulators of a motor protein which has microtubule stimulated ATPase	XX
CC	immunomodulators and antiinflammatory. Accordingly, through gene	XX	activity, a method of testing for ATPase activity of microtubule motor	XX
CC	therapy, they can be used for the treatment of cancer, hyperplasias,	XX	proteins, methods to identify candidate agents that bind to a target	XX
CC	restenosis, cardiac hypertrophy, immune disorders and inflammation. This	XX	protein or act as a modulator of the binding characteristics or	XX
CC	polypeptide sequence is human KID protein (Seqid 6) of the invention.	XX	biological activity of a target protein, modulators of the target	XX
XX	Sequence 346 AA;	XX	protein, and methods of treating cellular proliferation disorders such as	XX
SQ		XX	cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders	XX
	Query Match 93.9%; Score 1763; DB 7; Length 346;	XX	and inflammation, for treating disorders associated with kinesin-like DNA	XX
	Best Local Similarity 100.0%; Pred. No. 1.8e-170;	XX	binding protein (KID) and for inhibiting KID. The sequences are used for	XX
	Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	screening for modulators of motor proteins useful for treating cellular	XX
Qy	26 GRCRLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85	Qy	26 GRCRLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85	Qy
Db	2 GRCRLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61	Db	2 GRCRLSKIGATRRPPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61	Db
Qy	86 LKYOQDAFYGERSQOQDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 145	Qy	86 LKYOQDAFYGERSQOQDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 145	Qy
Db	62 LKYOQDAFYGERSQOQDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 121	Db	62 LKYOQDAFYGERSQOQDIYAGSVQPIRLHLEGGQNASVLAYGPTGAGKTHMLGSPQPG 121	Db
Qy	146 VIPALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205	Qy	146 VIPALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 205	Qy
Db	122 VIPALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181	Db	122 VIPALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRGN 181	Db
Qy	206 LIPGLSQKPISSFADFERRHFLPASRRTVGATRLNQRSSRSHAVLLVKVDQRELRAPFRQ 265	Qy	206 LIPGLSQKPISSFADFERRHFLPASRRTVGATRLNQRSSRSHAVLLVKVDQRELRAPFRQ 265	Qy
Db	182 LIPGLSQKPISSFADFERRHFLPASRRTVGATRLNQRSSRSHAVLLVKVDQRELRAPFRQ 241	Db	182 LIPGLSQKPISSFADFERRHFLPASRRTVGATRLNQRSSRSHAVLLVKVDQRELRAPFRQ 241	Db
Qy	266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGRLPRVYRDSKL 325	Qy	266 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGRLPRVYRDSKL 325	Qy
Db	242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGRLPRVYRDSKL 301	Db	242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSFLVGLKVVDALNQGRLPRVYRDSKL 301	Db
Qy	326 TRLLQDSLGGSAHSILIANIAPERRFYDITVSALNFAARSKEVIN 370	Qy	326 TRLLQDSLGGSAHSILIANIAPERRFYDITVSALNFAARSKEVIN 370	Qy
Db	302 TRLLQDSLGGSAHSILIANIAPERRFYDITVSALNFAARSKEVIN 346	Db	302 TRLLQDSLGGSAHSILIANIAPERRFYDITVSALNFAARSKEVIN 346	Db
RESULT 8				
ADQ60232				
ID	ADQ60232 standard; protein; 346 AA.			
XX				
AC	ADQ60232;			
XX				
DT	23-SEP-2004 (first entry)			
XX				
DE	Human microtubule motor protein #3.			
XX				
KW	Human; microtubule motor protein; cellular proliferation disorder;			
KW	cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;			
KW	inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;			
KW	arthritis; graft rejection; inflammatory bowel disease; angioplasty.			

```
QY 326 TRLLQSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370
Db 302 TRLLQSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 9
AAB56650
ID AAB56650 standard; protein; 460 AA.
XX AC
XX AAB56650;
XX 13-MAR-2001 (first entry)
XX DE
XX Human prostate cancer antigen protein sequence SEQ ID NO:1228.
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX OS
XX Homo sapiens.
XX FN
XX WO200055174-A1.
XX PD
XX 21-SEP-2000.
XX XX
XX 08-MAR-2000; 2000WO-US005988.
XX PF
XX 12-MAR-1999; 99US-0124270P.
XX XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX (ROSE/) ROSEN C A.
XX XX
XX Rosen CA, Ruben SM;
XX PI
XX WPI; 2000-587513/55.
XX DR
XX N-PSDB; AAF15853.
XX XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer
XX PT antigens, useful for treatment, prevention, and diagnosis of disorders
XX PT such as prostate cancer.
XX XX
XX Claim 11; Page 1649-1651; 2338pp; English.
XX XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention
XX SQ
XX Sequence 460 AA;
XX Query Match 93.9%; Score 1763; DB 3; Length 460;
XX Best Local Similarity 99.4%; Pred. No. 2.8e-170;
XX Matches 345; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 24 GAGCRLSKIGATRRPPPARVAVLRPFVDTAGADPPCVRGWDSCSLFIANWRHQ 83
Db 2 GAGCRLSKIGATRRPPPARVAVLRPFVDTAGADPPCVRGWDSCSLFIANWRHQ 61
QY 84 EFLTKYQDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPEQ 143
|||||
```

26	GRCLSKIGATRRPPPARVRAVRLRPFDVGTAGASDPPCVRGMDSCSLEIATANWRNHQET	85
2	GRCLSKIGATRRPPPARVRAVRLRPFDVGTAGASDPPCVRGMDSCSLEIATANWRNHQET	61
86	LKYQDFAPYGERSTQODIYAGSVQPIRLRHLLGQNASVLA YCPTGAGKTHTMLGSPQPG	145
62	LKYQDFAPYGERSTQODIYAGSVQPIRLRHLLGQNASVLA YCPTGAGKTHTMLGSPQPG	121
146	VIPRALMDLLQLTREEGAERCPWALSVTWMSYLEIYOEKVLDLLDPASGDVLVIREDCRGN	205
122	VIPRALMDLLQLTREEGAERCPWALSVTWMSYLEIYOEKVLDLLDPASGDVLVIREDCRGN	181
206	LIPGLSOKPTSSPADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPRQ	265
182	LIPGLSOKPTSSPADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVDQERLAPRQ	241
266	REGKLYLIDLAGSEDNRRTGNKGLRUKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL	325
242	REGKLYLIDLAGSEDNRRTGNKGLRUKESGAINTSFLVLGKVVDALNOGLPRVPYRDSKL	301
326	TRLLQDSJGGSAAHILIANIAPERRFYLDVTVSALNFAARKEVIN	370
302	TRLLQDSJGGSAAHILIANIAPERRFYLDVTVSALNFAARKEVIN	346

[illegible]

XX	13-JUL-2004.	
XX	06-MAR-2002; 2002US-00093317.	
XX	20-APR-1999; 99US-00295612.	
XX	20-JUN-2000; 2000US-00597292.	
XX	28-NOV-2000; 2000US-00724224.	
XX	(CYTO-) CYTOKINETICS INC.	
XX	Beraud C;	
XX	WPI; 2004-532491/51.	
XX	N-PSDB; ADQ60233.	
XX	New isolated microtubule motor protein, useful for screening modulators	
XX	for treating cellular proliferation disorders such as cancer,	
XX	hyperplasia, restenosis, cardiac hypertrophy, immune disorders and	
XX	inflammation.	
XX	Claim 1; SEQ ID NO 8; 26pp; English.	
XX	The invention relates to human microtubule motor proteins and the nucleic	
XX	acid sequences thereof.	



cell proliferative disorder; cancer; atherosclerosis; viral infection;  
HIV; neurological disorder; Parkinson's disease; Alzheimer's disease;  
stroke; human.

Homo sapiens.

Key Location/Qualifiers

9  
Domain /note = kinesin motor catalytic domain

11..25  
Domain /note = kinesin motor domain

22  
Modified-site /note = "potential phosphorylation site"

62..77  
Domain /note = kinesin motor domain

63..647  
Domain /note = kinesin motor domain

65  
Modified-site /note = "potential phosphorylation site"

84  
Modified-site /note = "potential phosphorylation site"

87  
Modified-site /note = "potential phosphorylation site"

95..116  
Domain /note = kinesin motor domain

97  
Modified-site /note = "potential phosphorylation site"

117  
Modified-site /note = "potential phosphorylation site"

122..135  
Domain /note = kinesin motor domain

127  
Modified-site /note = "potential phosphorylation site"

138  
Modified-site /note = "potential phosphorylation site"

144..162  
Domain /note = kinesin motor domain

149  
Modified-site /note = "potential phosphorylation site"

202..226  
Domain /note = kinesin motor domain

214  
Modified-site /note = "potential phosphorylation site"

217  
Modified-site /note = "potential phosphorylation site"

227  
Modified-site /note = "potential phosphorylation site"

248..287  
Domain /note = kinesin motor domain

254  
Modified-site /note = "potential phosphorylation site"

274  
Modified-site /note = "potential phosphorylation site"

288  
Modified-site /note = "potential phosphorylation site"

292  
Modified-site /note = "potential phosphorylation site"

295..325  
Domain /note = kinesin motor domain

328  
Modified-site /note = "potential phosphorylation site"

356  
Modified-site /note = "potential phosphorylation site"

386  
Modified-site /note = "potential phosphorylation site"

397..399  
Region /note = "leucine zipper pattern"

454  
Modified-site /note = "potential phosphorylation site"

469  
Modified-site /note = "potential phosphorylation site"

489  
Modified-site /note = "potential phosphorylation site"

495  
Modified-site /note = "potential phosphorylation site"

518..539  
Region /note = "leucine zipper pattern"

525..546  
Region /note = "leucine zipper pattern"

605  
Modified-site /note = "potential phosphorylation site"

649  
Modified-site /note = "potential phosphorylation site"

678  
Modified-site /note = "potential phosphorylation site"

682  
Modified-site /note = "potential phosphorylation site"

703  
Modified-site /note = "potential phosphorylation site"

753  
Modified-site /note = "potential phosphorylation site"

760  
Modified-site /note = "potential glycosylation site"

767  
Modified-site /note = "potential phosphorylation site"

790  
Modified-site /note = "potential phosphorylation site"

805  
Modified-site /note = "potential phosphorylation site"

814  
Modified-site /note = "potential phosphorylation site"

816  
Modified-site /note = "potential phosphorylation site"

845  
Modified-site /note = "potential glycosylation site"

849  
Modified-site /note = "potential phosphorylation site"

893  
Modified-site /note = "potential phosphorylation site"

973  
Modified-site /note = "potential phosphorylation site"

977  
Modified-site /note = "potential phosphorylation site"

WO2004029205-A2.

08-APR-2004.

24-SEP-2003; 2003WO-US030198.

27-SEP-2002; 2002US-0414227P.

18-NOV-2002; 2002US-0427594P.

07-FEB-2003; 2003US-0445724P.

07-MAR-2003; 2003US-0453277P.

(INCY-) INCYTE CORP.

Swarnakar A, Ho A, Hafalia AJA, Emerling BM, Ison CH, Nguyen DB;  
Chien D, Lu DAM, Yang YG, Yue H, Murage J, Griffin JA, Burrill JD;  
Wang JT, Marquis JP, Blake JJ, Chawla NK, Lee S, Becha SD, Lee SY;  
Richardson TW, Bhatia UG, Zheng W, Sprague WW, Tang YT, Azimzai Y;  
WPI; 2004-305155/28.  
N-PSDB; ADO44181.

New human structural and cytoskeleton-associated proteins and  
polynucleotides for diagnosing, preventing or treating diseases  
associated with aberrant protein expression, e.g. cancer,  
atherosclerosis, HIV or stroke.

Claim 1; Page 150-152; 205pp; English.

The present sequence represents a structural and cytoskeleton-associated  
polypeptide (SCAP). The SCAP polypeptides and polynucleotides of the  
invention are useful for diagnosing, preventing or treating diseases or  
conditions associated with aberrant expression of SCAP, such as cell



Db 304 ---HIPYNSKUTRLLKDSLGNCQTIMIAAVSPSSVFYDDTYNTLKYANRAKDI 355

Search completed: November 5, 2004, 18:44:37  
Job time : 78.9417 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:44:52 ; Search time 86.8758 Seconds  
(without alignments)  
2081.726 Million cell updates/sec

Title: US-10-797-893-4

Perfect score: 2589

Sequence: 1 MPAAGSTQQRREMAAASA.....LEAKVLAQAEEKNHCFPTM 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	100.0	512	16	US-10-797-893-4
2	2467	95.3	487	16	US-10-797-893-8
3	2419	93.4	490	15	US-10-334-143-80
4	2151	83.1	460	9	US-09-925-300-1228
5	1873	72.3	370	16	US-10-797-893-2
6	1763	68.1	346	16	US-10-797-893-6
7	615.5	23.8	589	17	US-10-425-115-320209
8	601	23.2	377	16	US-10-437-963-159031
9	579	22.4	776	16	US-10-437-963-103943
10	578.5	22.3	783	17	US-10-425-115-332919
11	578.5	22.3	854	15	US-10-425-114-59708
12	578.5	22.3	1237	15	US-10-334-143-33
13	574.5	22.2	1232	14	US-10-116-712-670

14	574.5	22.2	1232	16	US-10-408-765A-2153	Sequence 2153, Ap
15	574.5	22.2	1235	15	US-10-334-143-8	Sequence 8, Appli
16	569	22.0	1030	15	US-10-425-114-62748	Sequence 62748, A
17	566	21.9	1001	17	US-10-425-115-231231	Sequence 231231,
18	564.5	21.8	548	15	US-10-108-260A-2692	Sequence 2692, Ap
19	561	21.7	1388	14	US-10-146-473-82	Sequence 82, Appl
20	561	21.7	1388	15	US-10-173-999-32	Sequence 32, Appl
21	561	21.7	1388	15	US-10-332-089-2	Sequence 2, Appli
22	561	21.7	1388	15	US-10-188-832-164	Sequence 164, App
23	560	21.6	992	17	US-10-739-930-7903	Sequence 7903, Ap
24	559.5	21.6	1401	15	US-10-287-226-142	Sequence 142, App
25	557.5	21.5	672	16	US-10-408-765A-1664	Sequence 1664, Ap
26	557.5	21.5	1232	14	US-10-116-712-664	Sequence 664, App
27	557.5	21.5	1232	14	US-10-116-712-669	Sequence 669, App
28	557	21.5	1324	15	US-10-287-226-314	Sequence 314, App
29	552.5	21.3	757	15	US-10-220-120-366	Sequence 366, App
30	552.5	21.3	757	16	US-10-363-829-316	Sequence 316, App
31	550	21.2	864	9	US-09-883-096-2	Sequence 1, Appli
32	546.5	21.1	1103	9	US-09-847-874A-1	Sequence 1, Appli
33	546.5	21.1	1103	14	US-10-458-162-1	Sequence 1477, Ap
34	546.5	21.1	1826	16	US-10-408-765A-1477	Sequence 306, App
35	545.5	21.1	834	15	US-10-287-226-306	Sequence 25, Appl
36	544.5	21.0	935	14	US-10-080-608A-25	Sequence 114, App
37	544.5	21.0	935	14	US-10-332-089-6	Sequence 6, Appli
38	543	21.0	409	15	US-10-370-685-114	Sequence 23, Appl
39	543	21.0	928	14	US-10-080-608A-23	Sequence 112, App
40	543	21.0	928	14	US-10-370-685-112	Sequence 176714,
41	541.5	20.9	1382	16	US-10-437-963-176714	Sequence 6, Appli
42	535	20.7	513	17	US-10-601-036-6	Sequence 2, Appli
43	533.5	20.6	1057	17	US-10-601-036-2	Sequence 472, App
44	532	20.5	1056	14	US-10-282-174-472	Sequence 474, App
45	532	20.5	1056	14	US-10-282-174-474	

#### ALIGNMENTS

#### RESULT 1

US-10-797-893-4  
; Sequence 4, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Human  
; US-10-797-893-4

Query Match	100.0%;	Score 2589;	DB 16;	Length 512;
Best Local Similarity	100.0%;	Pred. No. 3.3e+209;		
Matches 512;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MPAAGSTQQRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVVRVAVLRPFVDGTAG	60
Db	1	MPAAGSTQQRREMAAASAAAI	SGAGRCRLSKIGATRRPPPARVVRVAVLRPFVDGTAG	60
Qy	61	ASDPPCVRGMDSCSLSEIANWRNHQETLKYQFDAPFYGERSTQDDIYAGSVQPIRLHLEGG	120	
Db	61	ASDPPCVRGMDSCSLSEIANWRNHQETLKYQFDAPFYGERSTQDDIYAGSVQPIRLHLEGG	120	
Qy	121	NASVLAVGPTGAGTKHTMLGSPQPGVIPRALMDLLQLTREEGAGRPPWALSVTMSYLEI	180	

```

Db 121 NASVLAYGPTGAGKTHMLGSPQPGVPRALMDLLQLTREAGECRPAWLSVTMSYLEI 180
QY 181 YQKVLDDLPASGDLVIREDCRGNLILPGLSQKPISSFADPERHFLPASRNRVTGATRL 240
Db 181 YQKVLDDLPASGDLVIREDCRGNLILPGLSQKPISSFADPERHFLPASRNRVTGATRL 240
QY 241 NQSSRSRAVLVKKVDQERLAPFROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINT 300
Db 241 NQSSRSRAVLVKKVDQERLAPFROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINT 300
QY 301 SLFVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVTVAL 360
Db 301 SLFVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVTVAL 360
QY 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPPEAKRGARPEEIEIGSPPEMAAP 420
Db 361 NFAARKEVINRPTNESLOPHALGPVKLSQKELLGPPPEAKRGARPEEIEIGSPPEMAAP 420
QY 421 ASASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWMLMKTVEEKD 480
Db 421 ASASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGAPLLSTPKRERWMLMKTVEEKD 480
QY 481 LEIERLKTQKLEAKMLAQKAEKENHCPTM 512
Db 481 LEIERLKTQKLEAKMLAQKAEKENHCPTM 512

RESULT 2
US-10-797-893-8
; Sequence 8, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/797,893
; PRIOR FILING DATE: 2004-03-09
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 09/597,292
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Human
US-10-797-893-8

Query Match 95.3%; Score 2467; DB 16; Length 487;
Best Local Similarity 100.0%; Pred. No. 5.9e-199;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GRCRLSKIGATRRPPPARVAVRLPPFVDGTAGSDPPCVRGMDSCSLEIANRNHQT 86
Db 2 GRCRLSKIGATRRPPPARVAVRLPPFVDGTAGSDPPCVRGMDSCSLEIANRNHQT 61
QY 87 LKQYDFAFYGERSTQDDIYAGSVQPIRLHLLBQNASVLAYGPTGAGKTHMLGSPQPG 146
Db 62 LKQYDFAFYGERSTQDDIYAGSVQPIRLHLLBQNASVLAYGPTGAGKTHMLGSPQPG 121
QY 147 VTPRALMDLLQLTREAGECRPAWLSVTMSYLEIYQKVLDDLPASGDLVIREDCRGN 206
Db 122 VTPRALMDLLQLTREAGECRPAWLSVTMSYLEIYQKVLDDLPASGDLVIREDCRGN 181
QY 207 LIPGLSQKPISSFADPERHFLPASRNRVTGATRLNQRSSRAVLVKKVDQERLAPF 266
Db 182 LIPGLSQKPISSFADPERHFLPASRNRVTGATRLNQRSSRAVLVKKVDQERLAPF 241
QY 267 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSKL 326

```

```

Db 242 REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVADALNOGLPRVPYRDSKL 301
QY 327 TLLQDSLGSAHSILIANIAPERRFYLDVTVALNFAARKEVINRPTNESLOPHALGP 386
Db 302 TLLQDSLGSAHSILIANIAPERRFYLDVTVALNFAARKEVINRPTNESLOPHALGP 361
QY 387 VKLSQKELLGPPPEAKRGARPEEIEIGSPPEMAAPASOKLSPLQKLSMDPAMLERLLS 446
Db 362 VKLSQKELLGPPPEAKRGARPEEIEIGSPPEMAAPASOKLSPLQKLSMDPAMLERLLS 421
QY 447 LDRLLASQSGAPLLSTPKRERWMLMKTVEEKDLEIERLKTQKLEAKMLAQKAEKE 506
Db 422 LDRLLASQSGAPLLSTPKRERWMLMKTVEEKDLEIERLKTQKLEAKMLAQKAEKE 481
QY 507 NHCPTM 512
Db 482 NHCPTM 487

RESULT 3
US-10-334-143-80
; Sequence 80, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VVACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; FILE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; PRIOR FILING DATE: 2002-12-31
; PRIOR FILING DATE: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-80

Query Match 93.4%; Score 2419; DB 15; Length 490;
Best Local Similarity 100.0%; Pred. No. 6.6e-195;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGGSTQQRREMAAASAAAISGAGRCRLSKIGATRRPPPARVAVRLPPFVDGTAGS 62
Db 10 AAGGSTQQRREMAAASAAAISGAGRCRLSKIGATRRPPPARVAVRLPPFVDGTAGS 69
QY 63 DPPCVRGMDSCSLEIANRNHQTLYKQYDFAFYGERSTQDDIYAGSVQPIRLHLLBQNA 122
Db 70 DPPCVRGMDSCSLEIANRNHQTLYKQYDFAFYGERSTQDDIYAGSVQPIRLHLLBQNA 129
QY 123 SVLAYGPTGAGKTHMLGSPQPGVTPRALMDLLQLTREAGECRPAWLSVTMSYLEIYQ 182
Db 130 SVLAYGPTGAGKTHMLGSPQPGVTPRALMDLLQLTREAGECRPAWLSVTMSYLEIYQ 189
QY 183 EKVLDLLDPASGDLVIREDCRGNILIPGLSQKPISSFADPERHFLPASRNRVTGATRLNQ 242
Db 190 EKVLDLLDPASGDLVIREDCRGNILIPGLSQKPISSFADPERHFLPASRNRVTGATRLNQ 249
QY 243 RSRSHAVLLVKKVDQERLAPFROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSL 302
Db 250 RSRSHAVLLVKKVDQERLAPFROREGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSL 309
QY 303 FVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVTVALNF 362
Db 310 FVLGKVVADALNOGLPRVPYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDVTVALNF 369
QY 363 AARKEVINRPTNESLOPHALGPVKLSQKELLGPPPEAKRGARPEEIEIGSPPEMAAPAS 422
Db 370 AARKEVINRPTNESLOPHALGPVKLSQKELLGPPPEAKRGARPEEIEIGSPPEMAAPAS 429

```

QY 423 ASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDL 481  
 Db 430 ASOKLSPLQKLSMDPAMLERLLSLDRLLASQSGQAPLLSTPKRERVMVLMKTVEEKDL 488

RESULT 4

US-09-925-300-1228  
 ; Sequence 1228, Application US/09925300  
 ; Patent No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben,  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1228  
 ; LENGTH: 460  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (75)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (147)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; NAME/KEY: SITE  
 ; LOCATION: (435)  
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 ; US-09-925-300-1228

Query Match 83.1%; Score 2151; DB 9; Length 460;  
 Best Local Similarity 95.9%; Pred. No. 2.3e-172; Indels 0; Gaps 0;  
 Matches 424; Conservative 1; Mismatches 17;  
 QY 25 GAGCRLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHQ 84  
 Db 2 GAGCRLSKIGATRRPPPARVAVLRPFVDTAGASDPPCVRGMDSCSLEIANRNHQ 61  
 QY 85 ETLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGQNASVLAQPTGAGKTHMLGSPEQ 144  
 Db 62 ETLKYQDFAFYGERSTOODIYAGSVQPIRLHLEGQNASVLAQPTGAGKTHMLGSPEQ 121  
 QY 145 PGVPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 204  
 Db 122 PGVPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRG 181  
 QY 205 NILIPGLSQPISFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPF 264  
 Db 182 NILIPGLSQPISFADFPERHFLPASNRRTVGATRLNQRSSRSHAVLLVKVDQERLAPF 241  
 QY 265 RQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTLSFLVGLKVDALNQLPRVYRDS 324  
 Db 242 RQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTLSFLVGLKVDALNQLPRVYRDS 301  
 QY 325 KLTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNLSLOPHAL 384  
 Db 302 KLTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVINRPTNLSLOPHAL 361  
 QY 385 GPVKLSQKELLGPPPEAKRGPPEEETGSPPEMAAPASQKLSPLQKLSMDPAMLERL 444  
 Db 362 GPVKLSQKELLGPPPEAKRGPPEEETGSPPEMAAPASQKLSPLQKLSMDPAMLERL 421  
 QY 445 LSLDRLLASQSGQAPLLSTPK 466  
 Db 422 LQLGPSACLPGEPPSPVEYFK 443

RESULT 5

US-10-797-893-2  
 ; Sequence 2, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 370  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-10-797-893-2

Query Match 72.3%; Score 1873; DB 16; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-149; Indels 0; Gaps 0;  
 Matches 369; Conservative 0; Mismatches 0;  
 QY 3 AAGSTQORRREMAAASAAISGAGRCRLSKIGATRRPPPARVAVLRPFVDTAGAS 62  
 Db 2 AAGSTQORRREMAAASAAISGAGRCRLSKIGATRRPPPARVAVLRPFVDTAGAS 61  
 QY 63 DPCVGRMDSCSLEIANRNHQTLYQDFAFYGERSTOODIYAGSVQPIRLHLEGQNA 122  
 Db 62 DPCVGRMDSCSLEIANRNHQTLYQDFAFYGERSTOODIYAGSVQPIRLHLEGQNA 121  
 QY 123 SVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQ 182  
 Db 122 SVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLQLTREAGAGRPWALSVTMSYLEIYQ 181  
 QY 183 EKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFPERHFLPASNRRTVGATRLN 242  
 Db 182 EKVLDLDPASGDLVIREDCRGNILIPGLSQKPISSFADFPERHFLPASNRRTVGATRLN 241  
 QY 243 RSSRSHAVLLVKVDQERLAPFQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTL 302  
 Db 242 RSSRSHAVLLVKVDQERLAPFQREGKLYLIDLAGSEDRRTGNKGLRLKESGAINSTL 301  
 QY 303 FVLGKVDALNQLPRVYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 362  
 Db 302 FVLGKVDALNQLPRVYRDSKLTLLQDSLGSAHSILIANIAPERRFYLDTVSALNF 361  
 QY 363 AARSKEVIN 371  
 Db 362 AARSKEVIN 370

RESULT 6

US-10-797-893-6  
 ; Sequence 6, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20



Db 58 EYKLDAPFGQSRVCEIFDOEVSAGVIGPIEGTNATVAYGATSGSKTYTMOCTEDLP 117  
 Qy 147 VIPRALMDLLQTLREGEAGRPWALSVMYSLEYIQBKVLDPDLDPAGDLVIREDCRNI 206  
 Db 118 LMXSTVALCTGT-----WC-SVEISYEVYMERCYDLEPKAREINVLDDKGNL 167  
 Qy 207 LIPGLOKPTSSPADFERHFLPASRNRTVGATRNQSSRSHAVLLVKVQORERLAPFRQ 266  
 Db 168 QKGLAWPVKRSLEEFHEIYISIGVQRKVAHTGLNDVSSRSHAVLSIRITT-----DV 220  
 Qy 267 REGKLYLI---DLAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVYRD 323  
 Db 221 VKGNLITXDLNAGSDNRRTGNKGLRKESGAINTSFLVGLKVVDALNOGLPRVYRD 280  
 Qy 324 SKTRLLQDSIGGSAHSILIANIAPERRFFLDYTSALNFAARSKEVINRNPFTNBSLOPHA 383  
 Db 281 SKLRLQDSIGGSAHSILIANIAPERRFFLDYTSALNFAARSKEVINRNPFTNBSLOPHA 383  
 Qy 384 LGPVKLSQKLLGPPPEAK-----RARGPEE-----EIGSGPEPMAAPASASQ 425  
 Db 329 SSASKKXKDKVDM---EAKLRAWLESKGTKSQRMDCGLLSPNAIKTPLSMNSQ 377

RESULT 9

US-10-437-963-103943  
 ; Sequence 103943, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 103943  
 ; LENGTH: 776  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_101324C.1.pap  
 US-10-437-963-103943

Query Match 22.4%; Score 579; DB 16; Length 776;  
 Best Local Similarity 30.6%; Pred. No. 1.6e-39;  
 Matches 162; Conservative 100; Mismatches 209; Indels 58; Gaps 12;  
 Qy 2 PAA-----GGSTQORREMAAASAAATSGAGRCRLSKIGATRPPPARVRVAVLRPPFVD 56  
 Db 173 PAARKLSLGGMAARLKTAGEAGNGDAAG-----SRINVFVLRPMSR 217  
 Qy 57 GTAGADPPPCVRGMDSCSLBIANRHNQETLK-----YQDFAFYGERSTQODIYAGS 108  
 Db 218 KEDAGSRSCVKIIVNKDVLTTFASSETDYLRLKVRGRHFCDFSPDPTTQAEVYSTT 277  
 Qy 109 VQPIRLHLLGQNASVLAYGPTCAGKTHMLGSPGQVPIPRALMDLLQTLTREGEAGRP 168  
 Db 278 TSDLVGVGLQNGTVFCYCATGAGKTYTMLGTMESPGVMVLAIKDLFTKVRQSRSDGNH 337  
 Qy 169 WALSVTMSYLEIYQEKVLDLDPASGLVIREDCRGNILIPGLSKPISGFADFERHFLP 228  
 Db 338 ---SIQLSYLEVNEYTRDLSFGR-PLLRREDKQTV-AAGLTHYRAYSTDEVMKLLQ 392  
 Qy 229 ASNRVTGATRNQSRSHAVLLVKVDQRE-RLAPFRQREGKLYLIDLAGSDNRRTGN 287  
 Db 393 GNQNRTEPTFRVNETSSRSHAILQIVIVEYSIDGGSIVTVRVGKLSLIDLAGSERALATDQ 452

Qy 288 KGLRKESGAINTSLFVLGKVVDALNOGLPRVYRDSKLTLLQDSIGGSAHSILIANIA 347  
 Db 453 RTQRSIEGANINSLALSSCINALVEGKHIPYRNKLTQLLQKDSLGSCNTVMIANIS 512  
 Qy 348 PERRFYLDYTSALNFAARSKEVINRNPFTNBSLOPHALGPVKLSQKELL-----GPPEAKR 402  
 Db 513 PSNLSFGETQNTLHWADRAKEIKTKALTANEEVLRTVDTSETDQAKLVLELQKENSELRQ 572  
 Qy 403 ARGPEEEIIGSPBPMAAPASASQKLSPLQKLSNDWPAWLERLLSLDRLLASQSGCAPLL 462  
 Db 573 QLARQQOKLLTVQAOTLASNASPOQSPASQAISTPCSTQKVRKRSILAGN-----CF 625  
 Qy 463 STPKREMV-----LMKTVEEKDLEIERLTKKKELEAKMLAQKAE 503  
 Db 626 NTFDSKRPAAENAQRDLQKRVKAMEAIEKMK-KEHLLQK-----QKDE 670

RESULT 10

US-10-425-115-332919  
 ; Sequence 332919, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 332919  
 ; LENGTH: 783  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(783)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_66737C.1.pap  
 US-10-425-115-332919

Query Match 22.3%; Score 578.5; DB 17; Length 783;  
 Best Local Similarity 32.9%; Pred. No. 1.8e-39;  
 Matches 167; Conservative 94; Mismatches 205; Indels 41; Gaps 13;

Qy 11 RRREMAAASAAATSGAGRCRLSKIGA-TRPPPARVRVAVLRPPFVDGTAGSDPPCVRG 69  
 Db 180 RKLUSLGGVGGKUKAAG-----EVGAGTAAESRILVFLRPMGRKEKAGSRSCVKI 234  
 Qy 70 MDSCSLBIANRHNQETLK-----YQDFAFYGERSTQODIYAGSVQPIRLHLLGQN 121  
 Db 235 VNRKEVLTESASENDYLRKGRDHSFDFSDVFPDSTQAEVYSTADLVGVLOGRN 294  
 Qy 122 ASVLAVGPTCAGKTHMLGSPGQVPIPRALMDLLQTLTREGEAGRPWALSVMYSLEYI 181  
 Db 295 GTVFCYCATGAGKTYTMLGTMENPGVMVLAINDLPSKVTQKNH-----SIKLSLEYI 347  
 Qy 182 QEKVLDLDPASGLVIREDCRGNILIPGLSKPISGFADFERHFLPASNRVTGATRN 241  
 Db 348 NETVRDLSLSPGS-PLNLRDQKQ-IVAAGLTQBSVYSTDEVMELLQKGNKRNTEPTFRV 405  
 Qy 242 QRSRSHAVLLVKVQORERLAPFRQREGKLYLIDLAGSDNRRTGNKGLRKESGAINTS 301  
 Db 406 ETSRSHAVLVQVVEYRSLDGVNVKAGKLSLIDLAGSERALATDQRTQRSIEGANINRS 465  
 Qy 302 LFLVGLKVVDALNOGLPRVYRDSKLTLLQDSIGGSAHSILIANIAPERRFYLDYTSALN 361  
 Db 466 LLALSSCINALVEGKHIPYRNKLTQLLQKDSIGGACNTVMIANISPSNLSFGETQNTLH 525



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 Search time 10.0875 Seconds  
(without alignments)  
3300.235 Million cell updates/sec

Title: US-10-797-893-6  
Perfect score: 1768  
Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARKEVIN 346

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1550	87.7	665	2 S62328	kinesin-like DNA b
2	594.5	33.6	664	2 T48258	kinesin-like prote
3	555.5	31.4	784	1 A55236	kinesin-related pr
4	550	31.1	703	2 A86319	F25116.11 protein
5	547	30.9	1254	2 T18277	kinesin heavy chai
6	536	30.3	784	2 T40594	probable kinesin-1
7	532	30.1	1226	2 I51617	kinesin-like prote
8	530.5	30.0	885	2 D86151	P22M8.8 protein -
9	527.5	29.8	1388	2 T30335	KLP2 protein - Afr
10	525	29.7	1070	2 T06733	kinesin homolog F2
11	524	29.6	1056	2 H84777	probable kinesin-r
12	523	29.6	1058	2 T47525	kinesin-related pr
13	521.5	29.5	1006	2 T02017	kinesin-related pr
14	515.5	29.2	1076	2 B84687	probable kinesin-1
15	514.5	29.1	699	1 S38982	kinesin-related pr
16	514	29.1	1066	1 A48669	kinesin-related pr
17	514	29.1	1695	2 A56921	kinesin family pro
18	513.5	29.0	701	1 B44259	kinesin-related pr
19	513.5	29.0	805	2 S64238	kinesin-related pr
20	513.5	29.0	1225	2 A56514	chromokinesin - ch
21	513	29.0	747	1 A57107	kinesin-related pr
22	511.5	28.9	1022	2 E84792	probable kinesin h
23	509.5	28.8	786	2 A53939	kinesin homolog KH
24	509.5	28.8	2954	2 T14156	kinesin-related pr
25	508	28.7	1229	2 T48959	kinesin-like prote
26	507.5	28.7	1231	2 A54803	microtubule-associ
27	507	28.7	928	2 T10164	kinesin heavy chai
28	504.5	28.5	330	2 B48935	kinesin-like prote
29	504	28.5	813	2 T46242	kinesin-like prote

30	501	28.3	1056	1 G02157	kinesin-like spind
31	500	28.3	1150	1 A55289	kinesin-like prote
32	498.5	28.2	935	2 T51930	kinesin (imported)
33	495.5	28.0	958	2 T20621	hypothetical prote
34	495	28.0	929	2 T51932	kinesin (imported)
35	491	27.8	963	1 A41919	kinesin heavy chai
36	490	27.7	742	1 S58691	kinesin-like heavy
37	488	27.6	987	2 T51360	kinesin-related pr
38	485.5	27.5	793	2 JC5831	kinesin-related pr
39	485.5	27.5	843	2 S44868	kinesin heavy chai
40	483	27.3	332	2 C48935	kinesin-like prote
41	482.5	27.3	883	2 T40128	kinesin-like prote
42	482.5	27.3	1584	1 JN0114	kinesin-related pr
43	482.5	27.3	1584	2 T15822	kinesin-like prote
44	482	27.3	1027	2 S37711	kinesin heavy chai
45	482	27.3	1031	1 A38713	kinesin heavy chai

ALIGNMENTS

RESULT 1

S62328  
kinesin-like DNA binding protein KID - human  
C;Species: Homo sapiens (man)  
C;Date: 23-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 16-Aug-2004  
C;Accession: S62328  
R;Tokai, N.; Fujimoto-Nishiyama, A.; Toyoshima, Y.; Yonemura, S.; Tsukita, S.; Inoue, J.; Inoue, J.  
EMBO J. 15; 457-467, 1996  
A;Title: Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and  
A;Reference number: S62328; MUID:96174806; PMID:8599929  
A;Accession: S62328  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-665 <OK>  
C;Superfamily: kinesin motor domain homology  
C;Keywords: ATP; nucleotide binding; P-loop  
F;44-374/Domain: kinesin motor domain homology <KMT>  
F;128-135/Region: nucleotide-binding motif A (P-loop)

Query Match 87.7%; Score 1550; DB 2; Length 665;  
Best Local Similarity 90.8%; Pred. No. 1.6e-117;  
Matches 314; Conservative 1; Mismatches 29; Indels 2; Gaps 2;

QY	2	GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET	61
DB	26	GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANRNHOET	85
QY	62	LKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGNAS-VLAYGPTGAGKTHMLGSPEQP	120
DB	86	LKYQDFAFYGERSTQDDIYAGSVQPIRLHLEGNQAKVILAYGPTGAGKT-THAQQPRAT	144
QY	121	GVTPRALMDLLQITREGAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN	180
DB	145	WGDPAGSHGCPAAHKGCGCRGPMGLSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGN	204
QY	181	ILIPGLSQPISSFADFPERHFLPASNRVTGATRLNQRSSRSRAVLAVKVDQERLAPER	240
DB	205	ILIPGLSQPISSFADFPERHFLPASNRVTGATRLNQRSSRSRAVLAVKVDQERLAPER	264
QY	241	QREGKLYLDLAGSENNRRTNKGLRKESGAINTSFLVLGKVVDALNQGPRVPPYRDSK	300
DB	265	QREGKLYLDLAGSENNRRTNKGLRKESGAINTSFLVLGKVVDALNQGPRVPPYRDSK	324
QY	301	LTRLLQDSLGSSAHSILIANIAPERFYLDTSALNFAARKEVIN	346
DB	325	LTRLLQDSLGSSAHSILIANIAPERFYLDTSALNFAARKEVIN	370

RESULT 2

T48258  
kinesin-like protein - Arabidopsis thaliana  
N;Alternate names: protein T1E22.130

A;Description: may be part of a motor protein that provides anterograde fast axonal transport.  
C;Superfamily: kinesin-related protein Kif3; kinesin motor domain homology  
C;Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop  
F;1-349/Domain: head globular #status predicted <HGL>  
F;20-350/Domain: kinesin motor domain homology <KMT>  
F;106-113/Region: nucleotide-binding motif A (p-loop)  
F;350-580/Domain: helical rod #status predicted <RD>  
F;581-784/Domain: tail globular #status predicted <TGL>  
F;112/Binding site: ATP (lys) #status predicted

Query Match	31.4%;	Score 555.5;	DB 1;	Length 784;
Best Local Similarity	39.5%;	Pred. No. 7.7e-37;		
Matches 140;	Conservative 56;	Mismatches 127;	Indels 31;	Gaps 10;
Qy	12	TRRP-----PPARYRVAVRLRPFDGTAGASDPPCV-----RGMDSCSLEIANW--RN	57	
Db	5	SRRPFGTSSQTPNECVQVVVRCRPMNRERSRSPVNVVYPNRGV-----VELQNVVDGN	60	
Qy	58	HQSTLKQYDFAYGERSTQDYYAGSVQVQILRHLLGQNASVLAYGPTGAGKTHTW----L	114	

Chinoeologs, A.; Ecker, J.R.; Palm, C.O.; Federspiel, N.A.; Kaul, S.; Wulce, O.; Alonso-  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86319  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-703 <STO>  
A;Cross-references: UNIPROT:Q9FZ77; GB:AE005172; NID:g9795601; PIDN:AAF98419.1; GSPDB:GK  
C;Genetics:  
A;Map position: 1

Query Match  
Best Local Similarity 31.1%; Score 550; DB 2; Length 703;  
38.6%; Pred. No. 1.8e-36;



Db 210 FPETTTQOEVYSTTGDVLEAVLGRNGSVFCYATGAGKTYTMTGMENFGVMVLAIKD 269  
QY 130 LLQLTREBAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGNILIPGLSQK 189  
Db 270 LFAKVRQSLDGNH---VVHLSYLEVNETVRDLSLSPGR-PLIUREDKQ-----VWALLQR 321  
QY 190 PISSFADPERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFR--OREGKLY 247  
Db 322 -----GNQNRTEPTRCNETSSRSRSHAILQVIVYKTRDASMMIISRVGKLS 367  
QY 248 LIDLAGEDNRRTGNKGLRKESGAINTSFLVGLKVDALNOGLPRVPYRDSKLTLLQD 307  
Db 368 LIDLAGEDNRRTGNKGLRKESGAINTSFLVGLKVDALNOGLPRVPYRDSKLTLLQD 307  
QY 308 SLGGSCTVMIANISPSQSFGETONTLHWADRAKEI 464  
Db 428 SLGGSCTVMIANISPSQSFGETONTLHWADRAKEI 464

## RESULT 5

T18277  
kinesin heavy chain - slime mold (Dictyostelium discoideum)  
C;Species: Dictyostelium discoideum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T18277  
R;deHostos, E.L.; McCaffrey, G.; Vale, R.D.  
submitted to the EMBL Data Library, November 1995  
A;Reference number: Z18953  
A;Accession: T18277  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1254 <DEH>  
A;Cross-references: UNIPROT:Q94463; EMBL:U41289; NID:g1526990; PID:g1526991; PIDN:AA077

Query Match 30.9%; Score 547; DB 2; Length 1254;  
Best Local Similarity 36.3%; Pred. No. 7.2e-36;  
Matches 122; Conservative 76; Mismatches 126; Indels 12; Gaps 5;

QY 15 PPARVAVRLRPFDVGTAGASPPCVRGWDSLSRI-ANWRNHQETLKYQPDAPFGR 73  
Db 24 PVSNIIVCVRLPTELEGRNHSIVHFDKSKSIRANGP-----QTFDRIFGY 77  
QY 74 STQDIYAGSVQPIRLHLLGQNASVLAYGPTGAGKTHMLGSPQPGVPRALMDL--- 130  
Db 78 ETQSQIPEDVAEPIVNDFLDGYHCTIAYGTAGSKTFTWGPDPDSHGIIIPVIESIFVG 137  
QY 131 LQLTREBAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGNILIPGLSQK 190  
Db 138 ISKMRKDT-SLSLAFCLKISALELYNEKLYDLYIASKNLNIREHKQGIYVEGISEIV 196  
QY 191 ISSFADPERHFLPASNRRTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQEGKLYLD 250  
Db 197 ITSIEEYFNLSNNRRASTKMSAASRSRSHVLMIELSQ-QNLSMWSKSKULFLVD 255  
QY 251 LAGSDNRRTGNKGLRKESGAINTSFLVGLKVDALNOGLPRVPYRDSKLTLLQDLSLG 310  
Db 256 LAGSDISHTKGAEGDRQAEAKNINLSLALGKVINALTGCVNYVPYRDSKLTTRVLQDLSLG 315  
QY 311 GSAHSILIANIAPERRRFLDTVSALNPAARSKEVIN 346  
Db 316 GNSKTSLIINCSPSNNEHEHTITLQFGTRAKTIKN 351

## RESULT 6

T40594  
probable kinesin-like protein - fission yeast (Schizosaccharomyces pombe)  
C;Species: Schizosaccharomyces pombe  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T40594; T39531  
R;Purnelle, B.; Goffeau, A.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, May 1998  
A;Reference number: Z21939

A;Accession: T40594  
A;Molecule type: DNA  
A;Residues: 1-784 <PUR>  
A;Cross-references: UNIPROT:O59751; EMBL:AL023587; PIDN:CAA19043.1; GSPDB:GN00067; SPDB:SPBC1685.15C  
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.  
submitted to the EMBL Data Library, July 1998  
A;Reference number: Z21861  
A;Accession: T39531  
A;Molecule type: DNA  
A;Residues: 484-784 <WOO>  
A;Cross-references: EMBL:AL031154; PIDN:CAA20063.1; GSPDB:GN00067; SPDB:SPBC1685.15C  
A;Experimental source: strain 972h-; cosmid c1685  
C;Genetics:  
A;Gene: SPDB:SPBC649.01c; SPDB:SPBC1685.15C  
A;Map position: 2  
A;Introns: 11/1

Query Match 30.3%; Score 536; DB 2; Length 784;  
Best Local Similarity 35.5%; Pred. No. 2.9e-35;  
Matches 139; Conservative 61; Mismatches 116; Indels 76; Gaps 12;

QY 20 VRVAVRLRPFDVGTAG--ASDPPCVRGWDSLSRIANWRNH----- 58  
Db 7 ISVAVRVRPPTEREKGLLAETPKSKEFLGDSLAVSNTSNTCTGTGIRKIVRVLDNDVL 66  
QY 59 -----OETL-----KYOFDAFYGERSTOODIYAGSVQPIRLHLLGQ 95  
Db 67 IFDPPENPLAKVQKSLLPAGKPRDVRYPADRLFGSEASQEDVYKGTTEPLDSVLQGY 126  
QY 96 NASVLAYGPTGAGKTHMLGSPQPGVI---PRALMDLLOLTREBAGRPWALSVTMSY 152  
Db 127 NATVFAYGATCGKTHITISRPDPDGIIFLTMRALD-----RVGLK-RTMNVVDISVY 180  
QY 153 LEIYQEKVLDL--DPASG-----LVIREDCRGNILIPGLSQKPISSFADPERHFLPASR 206  
Db 181 LEIYNEKIRDLLVQDPLSMKPKSLNICEAEQNVSVPLGLSYFTPTNLEEVMEIIRGNS 240  
QY 207 NRTVGATRLNQRSSRSRSHAVLLVKVDORERLAPFRQEGK-----LYLIDLAGED 256  
Db 241 NRTWSPTEANAVSRSHAVLQIYITQPKSGE-KQESSESONSHKVSFSDLAGSER 299  
QY 257 NRTGNKGLRKESGAINTSFLVGLKVDALNOGLPR---VPYRDSKLTLLQDLSLGS 312  
Db 300 ASATKRGKELVEGANINRSLLALGNCINSLSCE--PRRQHVFPYRDSKLTLLKFSLGGN 357  
QY 313 AHSILIANIAPERRRFLDTVSALNPAARSKEV 344  
Db 358 CRTCMIVCISPSSSEHYDETHNTLKYGNRAKNI 389

## RESULT 7

I51617  
kinesin-like protein 1 - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Aug-2004  
C;Accession: I51617; A48835; S48837  
R;Vernos, I.; Raats, J.; Hirano, T.; Heasman, J.; Karsenti, E.; Wylie, C.  
Cell 81, 117-127, 1995  
A;Title: Xklpi, a chromosomal Xenopus kinesin-like protein essential for spindle organi  
A;Reference number: A56221; MUID:95236444; PMID:7720067  
A;Accession: I51617  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1226 <VER>  
A;Cross-references: UNIPROT:Q91784; EMBL:X82012; NID:g562792; PIDN:CAA57539.1; PID:g562  
R;Vernos, I.; Heasman, J.; Wylie, C.  
Dev. Biol. 157, 232-239, 1993  
A;Title: Multiple kinesin-like transcripts in Xenopus oocytes.  
A;Reference number: A48835; MUID:93246085; PMID:8482413  
A;Accession: A48835  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA

	Query Match	30.0%;	Score 530.5;	DB 2;	Length 885;
	Best Local Similarity	36.3%;	Pred. No. 9.6e-35;		
	Matches 135;	Conservative	66;	Mismatches 133;	Indels 39; Gaps 9
QY	3	RCKLSKIGATRPPPARVRVAVLRP-	-FVDGTAGADPPCVR	---GMDSCSELTANWRN	57
	:				
Db	44	RSSASIGAANDNGVGRVRVAVLRPNRADESVAADADFCDVQLFKRLKLKRN--	--	101	
	:				
OY	58	HGFETIKYOFDAFYGSRSTOODIYAGSVOPILRHLLLEGONASVLAYGPTCAGKTHTM--	--LG	115	
	:				

## RESULT 10

Matches	131;	Conservative	67;	Mismatches	118;	Indels	44;	Gaps	9
---------	------	--------------	-----	------------	------	--------	-----	------	---

  

QY	20	VRVAVRLRPV	VDGTAGASDPP	CVRGMDSC	SLSEIANWRN	----	HOETLKYQ	FD	AFYGERST	75
Db	13	VQVLLRCRPF	SDDL-RSNA	QVLTNDL	QREVAVS	QNIACKHIDRV	-FTPD	KV	FGPSAQ	70
QY	76	QODIYAGS	VOPIIRLH	LEGQNASV	LVAGPT	CAGKTHWLG	-----	PPQ	PGVI	123
Db	71	QKQLYDAQ	AVPIVNEV	LEGFNCTI	FAYGQT	GTGKTYTMEG	ECRRKS	APCGGL	PAEAGVI	130
QY	124	PRALMDLL	QLTREB	AGBGPW	ALSVTMS	YLEIYOEK	VLDDLDP	-----	ASGD	172
Db	131	PRAVKQIF	-----	DTLEGQ	AEVSVK	VTFLYNEE	ITDLLAP	EDLSR	VAEEKQK	185
QY	173	IREDCRGN	LILPGL	SQKPISS	----	FADFERHF	YLPAS	RNRVTG	ATRLNQR	228
Db	186	LMEDGKGV	LVRLGLE	EIVTSANE	IFTLLER	---GSS	KRRTAE	TFLN	KQSSRS	241
QY	229	KVDQRE	LAPFRQ	--REG	KLYLID	LAGEDNR	RRTNGK	LRKLKES	GAINTSL	286
Db	242	THIKCATP	EGEBELI	KCGKLN	LVLAGSE	NIISRG	ARDGR	AREAGE	INKSLT	301
QY	287	LNQGLPRV	PYRDSK	LTRLLO	QSLGGS	AHSLIANI	APIERRE	FYDVT	VSALNFA	346
Db	302	LVEHLGHV	PYRDSK	LTRLRL	DSLGRR	TKTCIIAT	VS	PAVH	CLBETL	361

  

RESULT	12
T47525	kinesin-related protein-like - Arabidopsis thaliana
N:	Alternate names: protein F16L2.60
C:	Species: Arabidopsis thaliana (mouse-ear cress)
C:	Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:	Accession: T47525
R:	Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseid, M.; Mewes, H.W.;
A:	Reference number: Z24468
A:	Accession: T47525
A:	Status: preliminary
A:	Molecule type: DNA
A:	Residues: 1-1058 <JOR>
A:	Cross-references: UNIPROT:Q9LZU5; EMBL:AL162459
A:	Experimental source: cultivar Columbia; BAC clone F16L2

Query Match	29.6%;	Score 523;	DB 2;	Length 1058;
Best Local Similarity	33.4%;	Pred. No. 5e-34;		
Matches 126;	Conservative	74;	Mismatches 127;	Indels 50; Gaps 8
QY	5	RLSKIGATRRPPAR---	VRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANNRNHOET	61
Db		: : :   :   :   :	: : :   :   :   :	
31	RSSESNTNRNDKEGVNVQILRCRPLSDPEARHTFPVI-----SCN-----ENREV	80		
QY	62	LKYQ-----FDAFYGRSTOODIYAGSVQPILRHLLRGNQASVLAYGPTGACK	109	
Db		: : :   :   :   :	: : :   :   :   :	
81	AATQSIAGKHIDRHFAFDKVFGPASQQKDLYDQAICPIVFLEVLEGYNCTIIYAYGGTGTGK	140		
QY	110	THTMLGS-----PEQPGVIPRALMDLLQLTRREGAERCPWALSVTMSLIEIQEKL	161	
Db		: : :   :   :   :	: : :   :   :   :	
141	TYTMGGARKNGEFPSDAGVIPRAVKQIFDLLEAQAE-----YSMKVTELELYNEIS	195		
QY	162	DLLDP-----ASGDVLIREDCKNLTILPGLSQKPISSFADFERHFTLPASNRRTVG	211	
Db		: : :   :   :   :	: : :   :   :   :	
196	DLLAPSETIKFVDEKSKKSIALMEDGKGSFVRGLEEEIVSTANEIYKILEKGSAKRRTA	255		
QY	212	ATFLNQRSSSHAVLLVKVDORLERLAFPRQ--REKKLYLIDLACSEDNRNRTCNKGLRIKE	269	
Db		: : :   :   :   :	: : :   :   :   :	
256	ETLLNKQSQRSHSGIFTHTIHKNTPEGEMIKCGKLNLDVLAGESENISRGAEGRARE	315		

QY 270 SGAINSTSLFVLGKVVADNQLPRVYRDSKLTLLRLDLSLGSASHILANTAPERRFYL 329  
 Db 316 AGEINKSLTULGRVINALVEHSGHIPYRDSKLTLLRLSLGKTKTCVIATISPSIHCL 375  
 QY 330 DTVSALNFAARKEVIN 346  
 Db 376 ETLSTLDYAHRAKNIK 392

RESULT 13  
 T02017  
 Kinesin-related protein TKRP125 - common tobacco  
 C:Species: Nicotiana tabacum (common tobacco)  
 C>Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T02017  
 R:Asada, T.; Kuriyama, R.; Shibaoka, H.  
 J. Cell Sci. 110, 179-189, 1997  
 A:Title: TKRP125, a kinesin-related protein involved in the centrosome-independent organ  
 A:Reference number: Z14490; MUID:97196959; PMID:9044048  
 A:Accession: T02017  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1006 <ASA>  
 A:Cross-references: UNIPROT:023826; EMBL:D83711; NID:g2582970; PIDN:BAA23159.1; PID:g258  
 C:Genetics:  
 A:Note: TKRP125  
 C:Superfamily: kinesin-related protein Egs; kinesin motor domain homology  
 F:10-361/Domain: kinesin motor domain homology <KMT>

Query Match 29.5%; Score 521.5; DB 2; Length 1006;  
 Best Local Similarity 35.6%; Pred. No. 6.2e-34;  
 Matches 128; Conservative 69; Mismatches 118; Indels 45; Gaps 9;

QY 20 VRVAVLRPFVDGTAGASDPPCVRGMD-----SCSLEIANWRNHQETLKYQDFAFYGRS 74  
 Db 10 VQVLLRCRPFSDNLRNAPQVTCNDYQREVAVSQNIAG--KHIDRI-FTEDKVGPSA 66  
 QY 75 TQDIYAGSVQPIRLHLEGQNASVLAYGPTGAGTKHTMLGS-----PEQGV 122  
 Db 67 QQRDLYDQAIPIVNEVLGECFTIPAYGQTGTGTYTMEGCKRSKGPNGELPQEA 126  
 QY 123 IPRALMDLQLTREGAEGRPWALSVTMSYLEIYQEKVLDLDP-----ASGDIV 172  
 Db 127 IPRAVKQVDTLESQAE-----YSVKVTFLELYNEETDILLAPEDLKVALEDROKKQLP 181  
 QY 173 IREDCRGNILIPGLSQKPISS-----PADPERHFLPASNRVTGATRLNQRSSRHAVLLV 228  
 Db 182 LMEDGKGVLRGLEEIVTSANEIPTLLER---GSAKRRTAETLLNKKQSSRSLSFSI 237  
 QY 229 KYDQERLAPFRQ--REGKLYLIDLAGSDNRTGNKGLRKESGAINTSLFVLGKVYDA 286  
 Db 238 TTHIKENTPEGBELIKCGKLNLDVLAGSENISRGAREGRAREGEINKSLTLGRVINA 297  
 QY 287 LNQGLPRVYRDSKLTLLRLDLSLGSASHILANTAPERRFYLDTVSALNFAARKEVIN 346  
 Db 298 LVEHLCHIPYRDSKLTLLRLDLSLGGTKTCIIATVSPAVHCLTEETSLDLYAHRAKNIK 357

RESULT 14  
 B84687  
 probable kinesin-like spindle protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: B84687  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: B84420; MUID:20083487; PMID:10617197  
 A:Accession: B84687  
 A>Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-1076 <STO>  
 A:Cross-references: UNIPROT:Q9STB3; GB:AE002093; NID:g4580395; PIDN:AAD24373.1; GSPDB:G  
 C:Genetics:  
 A:Gene: At2G28620  
 A:Map position: 2  
 C:Superfamily: kinesin-related protein KIP1; kinesin motor domain homology

Query Match 29.2%; Score 515.5; DB 2; Length 1076;  
 Best Local Similarity 35.0%; Pred. No. 2.1e-33;  
 Matches 122; Conservative 68; Mismatches 132; Indels 27; Gaps 6;

QY 20 VRVAVLRPFVDGTAGASDPPCVRGMDSCSLEIA---NWRNHQETLKYQDFAFYGERSTQ 76  
 Db 51 IQVIVRCRPF--NSETRIQTFAVLTCNRKKEVAQVIAQKIDKTLFLDKVFGPTSOQ 109  
 QY 77 QDIYAGSVQPIRLHLEGQNASVLAYGPTGAGTKHTMLGS-----PEQGVIPRALM 128  
 Db 110 KDLYHQAVSPIVFEVLGYNCTIFAYGQTGTGTYTMEGARKKNGEIPSDAGVIPRAVK 169  
 QY 129 DLLQLTREGAEGRPWALSVTMSYLEIYQEKVLDL-----DPASGDIVIREDCRG 179  
 Db 170 QIFDILEAQSA---EYSLKVSFLELYNEELTDLLAPEETKFADDKSKKPLALMEDGKG 225  
 QY 180 NILIPGLSQKPISSPADPERHFLPASNRVTGATRLNQRSSRHAVLLKVDORERLAPF 239  
 Db 226 GVFRGLEEEIVSTADEIYKVKESAKRRTAETLLNKKQSSRSHSIFSITHIKECTPEG 285  
 QY 240 RQ--REGKLYLIDLAGSDNRTGNKGLRKESGAINTSLFVLGKVVDALMQGLPRVPYR 297  
 Db 286 EBIKVGKLNLDVLAGSENISRGAREGRAREGEINKSLTLGRVINALVEHSGHIPYR 345  
 QY 298 DSKLTLLRLDLSLGSASHILANTAPERRFYLDTVSALNFAARKEVIN 346  
 Db 346 ESKLTLLRLDLSLGGTKTCVIATVSPVHCLTEETSLDLYAHRAKHKN 394

RESULT 15  
 S38982  
 Kinesin-related protein KRP85 - sea urchin (Strongylocentrotus purpuratus)  
 N:Alternate names: kinesin-2 chain A; KRP (85/95) 85k chain  
 C:Species: Strongylocentrotus purpuratus (purple urchin)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S38982; S72551  
 R:Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.  
 Nature 366, 268-270, 1993  
 A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.  
 A:Reference number: S38982; MUID:94050179; PMID:8232586  
 A:Accession: S38982  
 A:Molecule type: mRNA  
 A:Residues: 1-699 <COL1>  
 A:Cross-references: UNIPROT:P46872; EMBL:L16993; NID:g295245; PIDN:AAA16098.1; PID:g2952  
 A:Accession: S72551  
 A:Molecule type: protein  
 A:Residues: 2-5,'X',7-11,59-64,125-132,222-226,'X',228-230 <COL2>  
 C:Complex: heterotrimer of a 115k chain and two kinesin-related chains of 95k (PIR:S5869  
 C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
 C:Keywords: ATP; heterotrimer; microtubule binding; nucleotide binding; P-loop  
 F:11-348/Domain: kinesin motor domain homology <KMT>  
 F:97-104/Region: nucleotide-binding motif A (P-loop)  
 F:103/Binding site: ATP (lys) #status predicted

Query Match 29.1%; Score 514.5; DB 1; Length 699;  
 Best Local Similarity 37.6%; Pred. No. 1.4e-33;  
 Matches 127; Conservative 61; Mismatches 135; Indels 15; Gaps 8;

QY 20 VRVAVLRPFVDGTAGASDPPCVRGMDSC--SLEIANWR--NHQETLKYQDFAFYGERST 75  
 Db 11 VRVVRCPRLSKETGQGFKSVK-MDEMRCVTQVTNPNASGEPPKSFDTFTVPAPGAK 69  
 QY 76 QDIYAGSVQPIRLHLEGQNASVLAYGPTGAGTKHTMLG---SPEQGVIPRALMDLLQ 132  
 Db 70 QTDVYNQATPIVDALIEGYNGTIPAYGQTGTGTYTMEGVRSQPELRGIIPNSFAHIFG 129

```

Qy 133 LTBEGAEGRPWALSVTMSYLEIYOEKVLDLL-DPASGDLVIREDCRGNILIPGLSQKPI 191
Db 130 HIAKEQENVR---FLVRVSYLEIYNEEVKDLGKQQHRLLEVKERPDVGVVVKDLSAFV 186
Qy 192 SSFADFERRHFLPASRNETVGATRLNQRSSRSHAVLLVKVDQERLAPFRQ--REGKLYLI 249
Db 187 NNADDMDRIMTLGNKNRSVGATNNWSSSRSHAIFTITLERSDMGLDKEQHVRVYKLUHV 246
Qy 250 DLAGEDNRRRTGNKGLRLKESCAINTSLFVLGKVVDALNOG-LPRVPYRDSKLTLLQDS 308
Db 247 DLASERQTKTGATGQRKATKINLSLSTLGNVISSLVGKSTHIYRNSKLTLLQDS 306
Qy 309 LGSASHILIANIAPERFYLDTVSALNFAARSKEVIN 346
Db 307 LGGNAKTVMCANIGFAEYNYDETISTLRYANRAKNIKN 344

```

Search completed: November 5, 2004, 18:45:33  
Job time : 11.0875 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2004, 18:34:12 ; Search time 71.0157 Seconds  
(without alignments)  
1747.786 Million cell updates/sec

Title: US-10-797-893-6

Perfect score: 1768

Sequence: 1 MGRCLSKIGATRRPPPARV.....FYLDTSALNFAARSKEVIN 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1768	100.0	346	7	Adc23342 Human kin
2	1768	100.0	346	8	Adq60232 Human mic
3	1768	100.0	487	7	Adc23344 Human kin
4	1768	100.0	487	8	Adq60234 Human mic
5	1763	99.7	370	7	Adc23338 Human kin
6	1763	99.7	370	8	Adq60228 Human mic
7	1763	99.7	490	7	Adk40973 Novel hum
8	1763	99.7	512	7	Adc23340 Human kin
9	1763	99.7	512	8	Adq60230 Human mic
10	1763	99.7	665	8	Adq9240 Human kin
11	1753	99.2	460	3	Aab56650 Human pro
12	559.5	31.6	784	4	Abb71112 Drosophil
13	559.5	31.6	548	7	Adm04007 Human pro
14	559.5	31.6	998	8	Ado44167 Structura
15	552.5	31.2	357	5	Aau76958 Novel hum
16	552.5	31.2	898	4	Abu53123 Intracell
17	552.5	31.2	898	5	Aau76957 Novel hum
18	552.5	31.2	898	7	Adc31082 Human nov
19	552.5	31.2	898	7	Adi15915 Human PP
20	551.5	31.2	603	6	Abu11606 Human MDD
21	549.9	31.1	905	5	Aau76967 Novel hum
22	535.5	30.3	677	4	Abb65183 Drosophil
23	532.5	30.1	408	8	Adn40550 Candida a
24	532.5	30.1	408	8	Adn40554 Candida a
25	532.5	30.1	972	8	Adn40548 Candida a

## ALIGNMENTS

RESULT 1

ADC23342

ID ADC23342 standard; protein; 346 AA.

XX AC ADC23342;

XX AC ADC23342;

DT 18-DEC-2003 (first entry)

DE Human kinesin-like DNA binding protein (KID) (SeqID 6).  
KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiac; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.

XX Homo sapiens.

XX US6387644-B1.

XX 14-MAY-2002.

XX 28-NOV-2000; 2000US-00724224.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI: 2003-706919/67.

XX N-PSDB; ADC23341.

PT Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.

XX Claim 1; SEQ ID NO 6; 26pp; English.

XX This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a

CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,  
 CC immunomodulators and antiinflammatory. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polypeptide sequence is human KID protein (SeqID 6) of the invention.

XX SQ Sequence 346 AA;

Query Match 100.0%; Score 1768; DB 7; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-173;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120  
 DB 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120  
 QY 121 GVIPALMDLLQLTREGAERPMWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180  
 DB 121 GVIPALMDLLQLTREGAERPMWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180  
 QY 181 ILIPGLSOKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPPR 240  
 DB 181 ILIPGLSOKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPPR 240  
 QY 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLSFLGKVVDALNQGLPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLSFLGKVVDALNQGLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 2

ID ADQ60232 standard; protein; 346 AA.

AC ADQ60232;

XX 23-SEP-2004 (first entry)

DE Human microtubule motor protein #3.

XX Human microtubule motor protein; cellular proliferation disorder;  
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX Homo sapiens.

XX US6762043-B1.

XX 13-JUL-2004.

XX 06-MAR-2002; 2002US-000933317.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX 28-NOV-2000; 2000US-00724224.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2004-532491/51.

XX N-PSDB; ADQ60231.

XX New isolated microtubule motor protein, useful for screening modulators

PT for treating cellular proliferation disorders such as cancer,

PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.

XX Claim 1; SEQ ID NO 6; 26pp; English.

XX The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase  
 CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention.

XX SQ Sequence 346 AA;

Query Match 100.0%; Score 1768; DB 8; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-173;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
 DB 1 MGRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQ 60  
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120  
 DB 61 TLKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYGPTGAGKTHMLGSPQP 120  
 QY 121 GVIPALMDLLQLTREGAERPMWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180  
 DB 121 GVIPALMDLLQLTREGAERPMWALSVMTMSYLEIQKVLDDLLDPASGDLVIREDCRGN 180  
 QY 181 ILIPGLSOKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPPR 240  
 DB 181 ILIPGLSOKPISSFADFERHFLPASRNRVTGATRLNQRSSSHAVLLVKVDQERLAPPR 240  
 QY 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLSFLGKVVDALNQGLPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGESEDNRRTGNKGLRKESGAINTSLSFLGKVVDALNQGLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 301 LTRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 3

ADQ23344

ID ADC23344 standard; protein; 487 AA.

XX AC ADC23344;

XX 18-DEC-2003 (first entry)

XX Human kinesin-like DNA binding protein (KID) (SeqID 8).

XX human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
 KW cytoskeletal; cardinals; immunomodulator; antiinflammatory; gene therapy;  
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
 KW cardiac hypertrophy; immune disorder; inflammation.

XX Homo sapiens.

XX OS



PN US6387644-B1.  
 XX 14-MAY-2002.  
 XX 28-NOV-2000; 2000US-00724224.  
 XX 20-APR-1999; 99US-00295612.  
 XX 20-JUN-2000; 2000US-00597292.  
 XX (CYTO-) CYTOKINETICS INC.  
 XX Beraud C;  
 XX WPI; 2003-706919/67.  
 XX N-PSDB; ADC23343.  
 XX Identifying a candidate agent as modulator of function of a target  
 PT protein for treating cellular proliferation disorders by adding a  
 PT candidate agent to a mixture of the target protein that  
 PT directly/indirectly produces ADP or phosphate.  
 XX Claim 1; SEQ ID NO 8; 26pp; English.  
 XX This invention relates to a novel method for high throughput screening  
 CC systems used to identify compounds for the treatment of cellular  
 CC proliferation disorders. Specifically, it refers to candidate agents that  
 CC are capable of modulating the activity of target proteins having motor  
 CC domains, such that the target protein directly or indirectly produces ADP  
 CC or phosphate. Furthermore, this activity can be determined using  
 CC fluorescence or absorbance readouts. The present invention describes a  
 CC method that identifies modulators of the target protein, which is a  
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals,  
 CC immunomodulators and antiinflammatory. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polypeptide sequence is human KID protein (SeqID 8) of the invention.  
 XX  
 SQ Sequence 487 AA;  
 Query Match 100.0%; Score 1768; DB 7; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-173;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWNHQE 60  
 Db 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWNHQE 60  
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIILRHLEGGQNASVLA YGPTGAGKTHTMLGSPEQ 120  
 Db 61 TLKYQDFAFYGERSTQDDIYAGSVQPIILRHLEGGQNASVLA YGPTGAGKTHTMLGSPEQ 120  
 QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 180  
 Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 180  
 QY 181 ILIPGLSQPISSFADPERHFLPASNRVTGATRLNQRSRSHAVLLVKVDORERLAPFR 240  
 Db 181 ILIPGLSQPISSFADPERHFLPASNRVTGATRLNQRSRSHAVLLVKVDORERLAPFR 240  
 QY 241 QREGKLYLIDLAGESSNRRGKGLKESGAINTSFLVLGKVVDALNQLPRVPYRDSK 300  
 Db 241 QREGKLYLIDLAGESSNRRGKGLKESGAINTSFLVLGKVVDALNQLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKVIN 346  
 Db 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDVTVSALNFAARSKVIN 346  
 RESULT 4  
 ADQ60234  
 ID ADQ60234 standard; protein; 487 AA.  
 XX  
 AC ADQ60234;

XX 23-SEP-2004 (first entry)  
 XX Human microtubule motor protein #4.  
 DE Human; microtubule motor protein; cellular proliferation disorder;  
 XX cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
 XX  
 OS Homo sapiens.  
 XX US6762043-B1.  
 XX 13-JUL-2004.  
 XX 06-MAR-2002; 2002US-00093317.  
 XX 20-APR-1999; 99US-00295612.  
 XX 20-JUN-2000; 2000US-00597292.  
 XX 28-NOV-2000; 2000US-00724224.  
 XX (CYTO-) CYTOKINETICS INC.  
 XX Beraud C;  
 XX WPI; 2004-532491/51.  
 XX N-PSDB; ADQ60233.  
 XX New isolated microtubule motor protein, useful for screening modulators  
 PT for treating cellular proliferation disorders such as cancer,  
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.  
 XX Claim 1; SEQ ID NO 8; 26pp; English.  
 XX The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase  
 CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention.  
 XX  
 SQ Sequence 487 AA;  
 Query Match 100.0%; Score 1768; DB 8; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-173;  
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWNHQE 60  
 Db 1 MGRCLSKIGATRRPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWNHQE 60  
 QY 61 TLKYQDFAFYGERSTQDDIYAGSVQPIILRHLEGGQNASVLA YGPTGAGKTHTMLGSPEQ 120  
 Db 61 TLKYQDFAFYGERSTQDDIYAGSVQPIILRHLEGGQNASVLA YGPTGAGKTHTMLGSPEQ 120  
 QY 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 180  
 Db 121 GVIPRALMDLLQLTREEGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 180

QY 181 ILIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240  
 DB 181 ILIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPFR 240  
 QY 241 QREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300  
 DB 241 QREGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSK 300  
 QY 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 301 LTRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346

RESULT 5  
 ADC23338

ID ADC23338 standard; protein; 370 AA.

XX AC ADC23338;

XX DT 18-DEC-2003 (first entry)

XX DE Human kinesin-like DNA binding protein (KID) (SeqID 2).

XX KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
 KW cytoskeletal; cardiac; immunomodulator; antiinflammatory; gene therapy;  
 KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
 KW cardiac hypertrophy; immune disorder; inflammation.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1 /note= "Encoded by ATGCA"

XX FT US6387644-B1.

XX PN 14-MAY-2002.

XX PD 28-NOV-2000; 2000US-00724224.

XX PF 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C;

XX PN WPI; 2003-706919/67.

XX DR N-PSDB; ADC23337.

XX PT Identifying a candidate agent as modulator of function of a target  
 PT protein for treating cellular proliferation disorders by adding a  
 PT candidate agent to a mixture of the target protein that  
 PT directly/indirectly produces ADP or phosphate.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC This invention relates to a novel method for high throughput screening  
 CC systems used to identify compounds for the treatment of cellular  
 CC proliferation disorders. Specifically, it refers to candidate agents that  
 CC are capable of modulating the activity of target proteins having motor  
 CC domains, such that the target protein directly or indirectly produces ADP  
 CC or phosphate. Furthermore, this activity can be determined using  
 CC fluorescence or absorbance readouts. The present invention describes a  
 CC method that identifies modulators of the target protein, which is a  
 CC kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiac,  
 CC immunomodulators and antiinflammatory. Accordingly, through gene  
 CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polypeptide sequence is human KID protein (SeqID 2) of the invention.  
 XX SQ Sequence 370 AA;

Query Match 99.7%; Score 1763; DB 7; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-173; Indels 0; Gaps 0;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 GRCLSKIGATRRPPPPARVRVAVRLRPFVDGTAGADPPPCVRGMDSCSLEIANRNHOET 61  
 DB 26 GRCLSKIGATRRPPPPARVRVAVRLRPFVDGTAGADPPPCVRGMDSCSLEIANRNHOET 85  
 QY 62 LKYOFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPPOPG 121  
 DB 86 LKYOFAFYGERSTQODIYAGSVQPIRLHLLLEGONASVLAAYGPTGAGKTHMLGSPPOPG 145  
 QY 122 VIPRALMDLLQLTREGEAGRPWALSVMTSYLEIQEKVLDDLLDPASGDLVIREDCRNI 181  
 DB 146 VIPRALMDLLQLTREGEAGRPWALSVMTSYLEIQEKVLDDLLDPASGDLVIREDCRNI 205  
 QY 182 LIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 241  
 DB 206 LIPGLSQPISSFADFERHFLPASNRVTGATRLNQRSSSRSHAVLLVKVDQERLAPFRQ 265  
 QY 242 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSKL 301  
 DB 266 REGKLYLIDLAGSDNRRTGNKGLRKESGAINTSLFVLGKVDALNOGLPRVPYRDSKL 325  
 QY 302 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 346  
 DB 326 TRLLQDSLGGSAHSILIANIAPERRFYLDTVSALNFAARKEVIN 370

RESULT 6

ADQ60228

ID ADQ60228 standard; protein; 370 AA.

XX AC ADQ60228;

XX DT 23-SEP-2004 (first entry)

XX DE Human microtubule motor protein #1.

XX KW Human; microtubule motor protein; cellular proliferation disorder;  
 KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
 KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
 KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.

XX OS Homo sapiens.

XX PN US6762043-B1.

XX PD 13-JUL-2004.

XX PF 06-MAR-2002; 2002US-00093317.

XX PR 20-APR-1999; 99US-00295612.

XX PR 20-JUN-2000; 2000US-00597292.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C;

XX WPI; 2004-532491/51.

XX PT New isolated microtubule motor protein, useful for screening modulators  
 PT for treating cellular proliferation disorders such as cancer,  
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.

XX PS Claim 1; SEQ ID NO 2; 26pp; English.

XX CC The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase

CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents a human microtubule motor protein  
 CC of the invention. Note: The specification states that this sequence is  
 CC encoded by the nucleic acid featured as SEQ ID NO:1, but this does not  
 CC appear to be the case.

XX  
 SQ Sequence 370 AA;

Query Match 99.7%; Score 1763; DB 8; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-173;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 DB 26 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85  
 QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 121  
 DB 86 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 145  
 QY 122 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 191  
 DB 146 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 205  
 QY 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQERLAPFRQ 241  
 DB 206 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQERLAPFRQ 265  
 QY 242 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 301  
 DB 266 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 325  
 QY 302 TRLLQDSLGSASHSILIANIAPERRFYLDITVSALNFAARSKVIN 346  
 DB 326 TRLLQDSLGSASHSILIANIAPERRFYLDITVSALNFAARSKVIN 370

RESULT 7

ADK40973

ID ADK40973 standard; protein; 490 AA.

XX  
 AC ADK40973;

XX  
 DT 06-MAY-2004 (first entry)

XX  
 DE Novel human kinase protein #80.

XX cytosolic; immunomodulator; cardiant; neuroprotective; nootropic;  
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
 KW cancer; peripheral nervous system; central nervous system;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; viral infection; prion infection;  
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
 KW attention disorder; cognition disorder; hypotension; hypertension;  
 KW psychotic disorder; neurological disorder; dyskinesia;  
 KW metabolic disorder; organ transplant rejection; enzyme.

OS Homo sapiens.

XX

PN WO2003057841-A2.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 31-DEC-2002; 2002WO-US041687.  
 XX  
 PR 31-DEC-2001; 2001US-0343169P.  
 XX  
 PA (GRIG/) GRIGORIEV I V.  
 PA (SUDA/) SUDARSANAM S.  
 XX  
 PI Grigoriev IV, Sudarsanam S;  
 XX  
 DR WPI; 2003-587115/55.  
 XX  
 DR New isolated, enriched or purified nucleic acid molecule encoding a  
 PT kinase polypeptide, useful for treating cancer, immune-related diseases,  
 PT cardiovascular disease, brain or neuronal-associated diseases and  
 PT metabolic disorders.  
 XX  
 PS Claim 1; SEQ ID NO 80; 491pp; English.

CC The invention relates to novel isolated, enriched or purified nucleic acid  
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule  
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and  
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the  
 CC polypeptide in (a), except that it lacks one or more, but not all, of an  
 CC N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-  
 CC terminal domain, a coiled-coil structure region, a spacer region and a C-  
 CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,  
 CC polypeptides, methods and substance are useful for treating cancers,  
 CC immune-related diseases or disorders, cardiovascular disease, brain or  
 CC neuronal-associated diseases, and metabolic disorders. The disorders are  
 CC preferably cancers of the tissues or of hematopoietic origin, diseases of  
 CC the central or peripheral nervous system, Alzheimer's disease,  
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,  
 CC viral infections, infections caused by prions, infections caused by  
 CC bacteria, infections caused by fungi, ocular diseases, migraines, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognition  
 CC disorders, hypotension, hypertension, psychotic disorders, neurological  
 CC disorders, dyskinesias, metabolic disorders and organ transplant  
 CC rejection. This sequence corresponds to one of the kinase polypeptides of  
 CC the invention.

XX  
 SQ Sequence 490 AA;

Query Match 99.7%; Score 1763; DB 7; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-173;  
 Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
 DB 34 GRCRLSKIGATRRPPPARVRVAVRLRPFVDGTAGASDPPCVRGMDSCSLEIANWRNHQET 93  
 QY 62 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 121  
 DB 94 LKYQDFAFYGERSTQODIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPQPG 153  
 QY 122 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 181  
 DB 154 VIPRALMDLLQTLTREGAEGRPWALSVTMSYLEIYQEKVLDLDPASGDLVIREDCRNI 213  
 QY 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQERLAPFRQ 241  
 DB 214 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQERLAPFRQ 273  
 QY 242 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 301  
 DB 274 REGKLYLIDLAGSEDRRTGNKGLRKESGAINTSFLVLGKVVDALNQGLPRVPYRDSKL 333  
 QY 302 TRLLQDSLGSASHSILIANIAPERRFYLDITVSALNFAARSKVIN 346

Db 334 TRLLQDSLGGSAHSILIANIAPERFVLDTVSALNFAARKEVIN 378

RESULT 8  
ADC23340  
ID ADC23340 standard; protein; 512 AA.  
XX  
AC ADC23340;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human kinesin-like DNA binding protein (KID) (SeqID 4).  
XX  
KW human; enzyme; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "Encoded by CA"  
FT  
XX  
FN US6387644-B1.  
XX  
PD 14-MAY-2002.  
XX  
PF 28-NOV-2000; 2000US-00724224.  
XX  
PR 20-APR-1999; 99US-00295612.  
PR 20-JUN-2000; 2000US-00597292.  
XX  
PA (CYTO-) CYTOKINETICS INC.  
XX  
PI Beraud C;  
XX  
DR WPI; 2003-706919/67.  
DR N-PSDB; ADC23339.  
XX  
PT Identifying a candidate agent as modulator of function of a target  
PT protein for treating cellular proliferation disorders by adding a  
PT candidate agent to a mixture of the target protein that  
PT directly/indirectly produces ADP or phosphate.  
XX  
PS Claim 1; SEQ ID NO 4; 26pp; English.  
XX  
CC This invention relates to a novel method for high throughput screening  
CC systems used to identify compounds for the treatment of cellular  
CC proliferation disorders. Specifically, it refers to candidate agents that  
CC are capable of modulating the activity of target proteins having motor  
CC domains, such that the target protein directly or indirectly produces ADP  
CC or phosphate. Furthermore, this activity can be determined using  
CC fluorescence or absorbance readouts. The present invention describes a  
CC method that identifies modulators of the target protein, which is a  
CC kinesin-like DNA binding protein (known as KID) as cytostatic, cardiants,  
CC immunomodulators and antiinflammatories. Accordingly, through gene  
CC therapy, they can be used for the treatment of cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
CC polypeptide sequence is human KID protein (SeqID 4) of the invention.  
XX  
SQ Sequence 512 AA;  
Query Match 99.7%; Score 1763; DB 7; Length 512;  
Best Local Similarity 100.0%; Pred. No. 7.4e-173;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GRCRLSKIGATRRPPARVAVRLRPFVDTGTAGADPPCVRGMDSCSLEIANWRHQET 61  
Db 27 GRCRLSKIGATRRPPARVAVRLRPFVDTGTAGADPPCVRGMDSCSLEIANWRHQET 86  
QY 62 LKQFDIFYGERSTQDDIYAGSVQPIIRHLLEGQNASVLAYGPTGAGKTHMLGSPQPG 121

Db 87 LKQFDIFYGERSTQDDIYAGSVQPIIRHLLEGQNASVLAYGPTGAGKTHMLGSPQPG 146

QY 122 VIPRALMDLLQLTREGAQRPAWLSVTMSYLIYQEKVLDLDPASGDLVIREDCRNI 181  
147 VIPRALMDLLQLTREGAQRPAWLSVTMSYLIYQEKVLDLDPASGDLVIREDCRNI 206

QY 182 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSRSHAVLLVKVQQRERLAPFRQ 241  
207 LIPGLSOKPISSPADFERHFLPASRNRVTGATRLNQRSSSRSHAVLLVKVQQRERLAPFRQ 266

QY 242 REGKLYLIDLAGSEDNRTGNKGLRKESGAINTSLFVLGKVVDALNOQLPRVYRDSKL 301  
267 REGKLYLIDLAGSEDNRTGNKGLRKESGAINTSLFVLGKVVDALNOQLPRVYRDSKL 326

QY 302 TRLLQDSLGGSAHSILIANIAPERFVLDTVSALNFAARKEVIN 346

Db 327 TRLLQDSLGGSAHSILIANIAPERFVLDTVSALNFAARKEVIN 371

RESULT 9  
ADQ60230  
ID ADQ60230 standard; protein; 512 AA.  
XX  
AC ADQ60230;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Human microtubule motor protein #2.  
XX  
KW Human; microtubule motor protein; cellular proliferation disorder;  
KW cancer; hyperplasia; restenosis; cardiac hypertrophy; immune disorder;  
KW inflammation; kinesin-like DNA binding protein; KID; autoimmune disease;  
KW arthritis; graft rejection; inflammatory bowel disease; angioplasty.  
XX  
OS Homo sapiens.  
XX  
PN US6762043-B1.  
XX  
PD 13-JUL-2004.  
XX  
PF 06-MAR-2002; 2002US-00093317.  
XX  
PR 20-APR-1999; 99US-00295612.  
PR 20-JUN-2000; 2000US-00597292.  
PR 28-NOV-2000; 2000US-00724224.  
XX  
PA (CYTO-) CYTOKINETICS INC.  
XX  
PI Beraud C;  
XX  
DR WPI; 2004-532491/51.  
XX  
PT New isolated microtubule motor protein, useful for screening modulators  
PT for treating cellular proliferation disorders such as cancer,  
PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
PT inflammation.  
XX  
PS Claim 1; SEQ ID NO 4; 26pp; English.  
XX  
CC The invention relates to human microtubule motor proteins and the nucleic  
CC acids encoding them. The invention also relates to a method of screening  
CC for modulators of a motor protein which has microtubule stimulated ATPase  
CC activity, a method of testing for ATPase activity of microtubule motor  
CC proteins, methods to identify candidate agents that bind to a target  
CC protein or act as a modulator of the binding characteristics of  
CC biological activity of a target protein, modulators of the target  
CC protein, and methods of treating cellular proliferation disorders such as  
CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
CC and inflammation, for treating disorders associated with kinesin-like DNA  
CC binding protein (KID) and for inhibiting KID. The sequences are used for  
CC screening for modulators of motor proteins useful for treating cellular  
CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
CC hypertrophy, immune disorders and inflammation, for treating disorders

CC associated with KID and for inhibiting KID and for treating autoimmune  
CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents a human microtubule motor protein  
CC of the invention. Note: The specification states that this sequence is  
CC encoded by the nucleic acid featured as SEQ ID NO:3, but this does not  
CC appear to be the case.

XX  
SQ Sequence 512 AA;

Query Match 99.7%; Score 1763; DB 8; Length 512;  
Best Local Similarity 100.0%; Pred. No. 7.4e-173; Indels 0; Gaps 0;  
Matches 345; Conservative 0; Mismatches 0;  
Qy 2 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61  
Db 27 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 86  
Qy 62 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121  
Db 87 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 146  
Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181  
Db 147 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 206  
Qy 182 LIPGLSQPISSPADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVQDRLAPFRQ 241  
Db 207 LIPGLSQPISSPADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVQDRLAPFRQ 266  
Qy 242 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDVALNOGLPRVPYRDSKL 301  
Db 267 REGKLYLDLAGSDNRRTGNKGLRKESGAINTSFLVLGKVDVALNOGLPRVPYRDSKL 326  
Qy 302 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
Db 327 TRLLQDSLGGSAAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 371

RESULT 10

ADQ09240  
ID ADQ09240 standard; protein; 665 AA.

XX AC ADQ09240;

XX DT 23-SEP-2004 (first entry)

XX DE Human KNSL4 protein SEQ ID NO:425.

XX KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;  
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;  
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;  
KW antiangiogenic; antiinflammatory; cardiovascular; cytosolic;  
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;  
KW human.

XX OS Homo sapiens.

XX PN W02004055050-A2.

XX PD 01-JUL-2004.

XX PF 10-DEC-2003; 2003MO-IB006434.

XX PR 10-DEC-2002; 2002US-0432699P.

XX PR 03-JUL-2003; 2003US-0485027P.

XX PA (ENDO-) ENDOCUBE SAS.

XX PA (CNRS ) CNRS CENT NAT RECH SCI.

XX PI Girard J, Amalric F, Roussigne M, Clouaire T;

XX DR WPI; 2004-525034/50.

DR N-PSDB; ADQ09241.

XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)  
PT responsive gene for preventing or treating e.g. cancer or inflammation,  
PT comprises modulating the interaction of a THAP polypeptide with a nucleic  
PT acid.

XX Example 47; SEQ ID NO 425; 612pp; English.

XX The present invention describes a method for modulating the expression of  
CC a thanatos (death)-associated protein (THAP) responsive gene. The method  
CC comprises modulating the interaction of a THAP-family polypeptide or its  
CC biological fragment with a nucleic acid, and so enhancing or repressing  
CC the expression of the THAP responsive gene. Also described: (1) a method  
CC of modulating the expression of a gene responsive to a THAP/chemokine  
CC complex; (2) a pharmaceutical composition comprising a THAP responsive  
CC element in a pharmaceutical carrier; (3) a transcription factor decoy  
CC consisting essentially of a THAP responsive element; (4) a cell  
CC comprising a transcription factor decoy described above; (5) methods of  
CC modulating the interaction between a nucleic acid and a THAP-family  
CC polypeptide or its biological fragment, or a nucleic acid and a  
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a  
CC cell comprising a viral vector which comprises a promoter operably linked  
CC to a nucleic acid encoding a THAP-family polypeptide or its biological  
CC fragment; (7) a method of constructing a cell which expresses a  
CC recombinant THAP-family polypeptide; (8) a method of ameliorating  
CC symptoms associated with a condition mediated by a THAP/chemokine complex  
CC; (9) methods of identifying a test compound that modulates transcription  
CC at a THAP responsive element or that modulates the transport of a  
CC chemokine into the nucleus; (10) methods for reducing the symptoms  
CC associated with a condition selected from excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative  
CC diseases; symptoms associated with a condition resulting from the  
CC activity of a chemokine or a THAP-family polypeptide in an individual; or  
CC symptoms associated with transcriptional repression or activation  
CC mediated by a THAP-family polypeptide in an individual; (11) a vector  
CC comprising a THAP responsive promoter operably linked to a nucleic acid  
CC encoding a detectable product; (12) a genetically engineered cell  
CC comprising the vector described above, or that expresses a THAP-family  
CC polypeptide or its biological fragment; (13) an in vitro transcription  
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,  
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-  
CC family polypeptide that does not bind to a chemokine. The pharmaceutical  
CC composition has antiangiogenic, antiinflammatory, cardiovascular,  
CC cytostatic, neuroprotective and osteopathic activities, and can be used  
CC as a THAP and THAP synthesis modulator. The composition can be used for  
CC modulating the expression of a THAP responsive gene. Modulation is useful  
CC for reducing symptoms of conditions such as excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative  
CC diseases. The present sequence is used in the exemplification of the  
XX present invention.

XX SQ Sequence 665 AA;

Query Match 99.7%; Score 1763; DB 8; Length 665;

Best Local Similarity 100.0%; Pred. No. 1.1e-172; Indels 0; Gaps 0;

Matches 345; Conservative 0; Mismatches 0;

Qy 2 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 61

Db 26 GRCLSKIGATRRPPPARVAVRLRPFDGTAGASDPPCVRGMDSCSLEIANWRNHQET 85

Qy 62 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 121

Db 86 LKYQDFAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLYAGPTGAGKTHMLGSPQPG 145

Qy 122 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 181

Db 146 VIPRALMDLLQLTREEGAGRPWALSVTMSYLEIQEKVLDLDPASGDLVIREDCRNI 205

Qy 182 LIPGLSQPISSPADFERHFLPASRNRVTGATRLNQRSSRSHAVLLVKVQDRLAPFRQ 241

Db 206 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQRLAPFRQ 265  
QY 242 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 301  
Db 266 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 325  
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
Db 326 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 370  
RESULT 11  
ID AAB56650 standard; protein; 460 AA.  
XX  
AC AAB56650;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Human prostate cancer antigen protein sequence SEQ ID NO:1228.  
XX  
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
KW wound; infectious disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200055174-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 08-MAR-2000; 2000WO-US005988.  
XX  
PR 12-MAR-1999; 99US-0124270P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX  
PI Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587513/55.  
DR N-PSDB; AAF15853.  
XX  
XX Prostate cancer associated gene sequences, referred to as prostate cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as prostate cancer.  
XX  
PS Claim 11; Page 1649-1651; 2338pp; English.  
XX  
XX AAF15566 to AAF16505 encode the human prostate cancer associated  
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
CC The prostate cancer antigens can have neuroprotective, cytostatic,  
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAB57303 represent sequences used in the exemplification of the present  
CC invention  
XX  
SQ Sequence 460 AA;

Query Match 99.2%; Score 1753; DB 3; Length 460;  
Best Local Similarity 99.4%; Pred. No. 6.8e-172;  
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLETANWRNHQET 61  
Db 4 GRCLSKIGATRRPPPARVAVRLRPFVDGTAGASDPPCVRGMDSCSLETANWRNHQET 63  
QY 62 LKYQFDIFYGERSSTQOQIYAGSVOPILRHLLLEGONASVLAYGPTGAGKTHTMLGSPQPG 121  
Db 64 LKYQFDIFYGEXSTQOQIYAGSVOPILRHLLLEGONASVLAYGPTGAGKTHTMLGSPQPG 123  
QY 122 VIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQKVLDDLDPASGDLVIREDCRNI 181  
Db 124 VIPRALMDLLQLTREGAEGRPWALSVTMSYLEIYQKVLDDLDPASGDLVIREDCRNI 183  
QY 182 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQRLAPFRQ 241  
Db 184 LIPGLSQKPISSFADFERHFLPASNRRTVGATRLNQRSSSHAVLLVKVDQRLAPFRQ 243  
QY 242 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 301  
Db 244 REGKLYLIDLAGSEDRRTGKGLRKESGAINTSFLVLGKVVVDALNOGLPRVPYRDSKL 303  
QY 302 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 346  
Db 304 TRLLQDSLGSAHSILIANIAPERRFYLDTVSALNFAARSKEVIN 348  
RESULT 12  
ID ABB71112 standard; protein; 784 AA.  
XX  
AC ABB71112;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 40128.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
DR N-PSDB; ABL15215.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Disclosure; SEQ ID NO 40128; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences



FT Modified-site /note = kinesin motor domain  
 FT 97 /note= "potential glycosylation site"  
 FT Modified-site 117  
 FT Modified-site 122  
 FT Domain 122..135  
 FT /note = kinesin motor domain  
 FT Modified-site 127  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 138  
 FT /note= "potential phosphorylation site"  
 FT Domain 144..162  
 FT /note = kinesin motor domain  
 FT Modified-site 149  
 FT /note= "potential phosphorylation site"  
 FT Domain 202..226  
 FT /note = kinesin motor domain  
 FT Modified-site 214  
 FT /note= "potential glycosylation site"  
 FT Modified-site 217  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 227  
 FT /note= "potential phosphorylation site"  
 FT Domain 248..287  
 FT /note = kinesin motor domain  
 FT Modified-site 254  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 274  
 FT /note= "potential glycosylation site"  
 FT Modified-site 288  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 292  
 FT /note= "potential phosphorylation site"  
 FT Domain 295..325  
 FT /note = kinesin motor domain  
 FT Modified-site 328  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 356  
 FT /note= "potential glycosylation site"  
 FT Modified-site 386  
 FT /note= "potential phosphorylation site"  
 FT Region 397..399  
 FT /note= "leucine zipper pattern"  
 FT Modified-site 454  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 469  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 489  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 495  
 FT /note= "potential phosphorylation site"  
 FT Region 518..539  
 FT /note= "leucine zipper pattern"  
 FT Region 525..546  
 FT /note= "leucine zipper pattern"  
 FT Modified-site 605  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 649  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 678  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 682  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 703  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 753  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 760  
 FT /note= "potential glycosylation site"  
 FT Modified-site 767  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 790  
 FT /note= "potential phosphorylation site"

FT Modified-site 805  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 814  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 816  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 845  
 FT /note= "potential glycosylation site"  
 FT Modified-site 849  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 893  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 973  
 FT /note= "potential phosphorylation site"  
 FT Modified-site 977  
 FT /note= "potential phosphorylation site"

WO2004029205-A2.

08-APR-2004.

24-SEP-2003; 2003WO-US030198.

27-SEP-2002; 2002US-0414227P.

18-NOV-2002; 2002US-0427594P.

07-FEB-2003; 2003US-0445724P.

07-MAR-2003; 2003US-0453277P.

(INCY-) INCYTE CORP.

Swarnakar A, Ho A, Hafalia AJA, Emerling BM, Ison CH, Nguyen DB;  
 Chien D, Lu DAM, Yang YG, Yue H, Murage J, Griffin JA, Burrill JD;  
 Wang JT, Marquis JP, Blake JJ, Chawla NK, Lee S, Becha SD, Lee SY;  
 Richardson TW, Bhatia UG, Zheng W, Sprague WW, Tang YT, Azimzai V;  
 WPI; 2004-305155/28.  
 N-PSDB; ABO44181.

New human structural and cytoskeleton-associated proteins and  
 polynucleotides for diagnosing, preventing or treating diseases  
 associated with aberrant protein expression, e.g. cancer,  
 atherosclerosis, HIV or stroke.

Claim 1; Page 150-152; 205pp; English.

The present sequence represents a structural and cytoskeleton-associated  
 polypeptide (SCAP). The SCAP polypeptides and polynucleotides of the  
 invention are useful for diagnosing, preventing or treating diseases or  
 conditions associated with aberrant expression of SCAP, such as cell  
 proliferative disorders (e.g. cancer or atherosclerosis), viral  
 infections (e.g. HIV) or neurological disorders (e.g. Parkinson's  
 disease, Alzheimer's disease or stroke). These may also be used for  
 assessing the effects of exogenous compounds on the expression of nucleic  
 acid and amino acid sequences of SCAP. The SCAP or its fragments are also  
 useful in screening compounds for effectiveness as agonist or antagonist  
 of the polypeptides, or in altering the expression of the target  
 polynucleotide and compounds that specifically bind to or modulate the  
 activity of the polypeptide.

Sequence 998 AA;

Query Match 31.6%; Score 559; DB 8; Length 998;

Best Local Similarity 38.9%; Pred. No. 7,1e-48;

Matches 132; Conservative 70; Mismatches 121; Indels 16; Gaps 7;

QY 19 RRVAVRLRPP--VDGTAGAS-----DPPCYRGWDSCLSEIANRNHQTLLK-YQFDAP 69

DB 11 QLMVALRVRLPISVALEGGATLIAHKVDQMVVLDPMDPDIIAHRRSKSLFDVA 70

QY 70 YGERSTOODIYAGSVQPIRLHLLLEGONASVLAAYGTGAKGTHMLGSPQPGVPRALMD 129

DB 71 FDFATQEMVYQATKSLIEGVISGNATVFAFGTGGCTYTLGTDQEPGIYVQTLD 130



QY 130 LLOUTREGAGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRGNILIPGSQX 189  
 Db 131 LFAIETSD--MEYEVMSYLEIYNEMIRDLNPSLGLYLELRDSKGVIVAGITEV 187  
 QY 190 PISSFADFERHFLPASNRVTGATRLNQRSSRSRSHAVLLVKVDQERLAPRQ--REGKLY 247  
 Db 188 STINAKIMQLMGKNGRQTEPAAQNTSSRSRSHAVLQVTVRQSRVKNILQEVROGRLF 247  
 QY 248 LIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDAL--NOGLPR-VPYRDSKLTREL 305  
 Db 248 MIDLAGSERASQTRNQRMKEGAHNRSLALGNCINALSDKSNKYINYRDSKLTREL 307  
 QY 306 QDSLGSSAHLIANIAPERFYLDTVSALNFAARSKEV 344  
 Db 308 KDSLGNSRTVMIAHISPASSAFESRNTLTLYAGRAKNI 346

RESULT 15

AAU76958

ID AAU76958 standard; protein; 357 AA.

XX AC AAU76958;

XX DT 21-MAY-2002 (first entry)

XX DE Novel human kinesin motor protein, HsKip3d motor domain.

XX KW Kinesin; motor protein; HsKip3d; cytoskeletal; immunosuppressive;  
 XX KW antiarthritic; antiinflammatory; vulnary; microtubule; inflammation;  
 XX KW cardiac hypertrophy; cellular proliferation disease; hyperplasia;  
 XX KW restenosis; cancer; solid tumour; carcinoma; autoimmune disease;  
 XX KW arthritis; graft rejection; inflammatory bowel disease; surgery;  
 XX KW angioplasty; nanotechnology; GeneChip array; expression monitoring;  
 XX KW human; enzyme.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Region 17..20

XX FT /note= "Interacts with the adenine ring of ATP to form  
 part of the nucleotide-binding site"

XX FT Region 113..120

XX FT /label= ATP/GTP binding\_site\_motif

XX FT Region 222..229

XX FT /label= SwitchII motif  
 XX FT /note= "Responsible for sensing of the gamma phosphate  
 and initiating the conformational change between ATP and  
 ADP bound forms of the enzyme"

XX FT Region 258..263

XX FT /label= SwitchII motif  
 XX FT /note= "Responsible for sensing of the gamma phosphate  
 and initiating the conformational change between ATP and  
 ADP bound forms of the enzyme"

XX PN W0200212268-A1.

XX PD 14-FEB-2002.

XX PF 03-AUG-2001; 2001WO-US024285.

XX PR 03-AUG-2000; 2000US-00632344.

XX PA (CYTO-) CYTOKINETICS INC.

XX PI Beraud C, Freedman R;

XX DR WPI; 2002-217176/27.

XX DR N-PSDB; ABK10348.

XX PT Novel microtubule motor protein for screening modulators of HsKip3d,  
 XX PT useful in treatment of hyperproliferative disease e.g. cancer, autoimmune  
 XX PT disease, arthritis, graft rejection, inflammatory bowel disease.

PS CC The invention describes an isolated novel human microtubule motor protein  
 XX CC (1). (1) is useful for screening modulators of HsKip3d. The method  
 CC comprises contacting (1) with a candidate agent in a test and control  
 CC concentration, and assaying for the level of HsKip3d activity, where the  
 CC HsKip3d activity is from binding activity or ATPase activity, and where a  
 CC change in activity between the test and control concentration indicates a  
 CC modulator. A composition containing (1) is useful in the treatment of e.g.  
 CC inflammation, cardiac hypertrophy, cellular proliferation diseases e.g.  
 CC hyperplasia, restenosis, cancer, including solid tumours of skin,  
 CC breast, brain, lung, heart, bone, genitourinary tract, liver, nervous  
 CC system, adrenal glands, haematologic, gynaecological and testicular  
 CC carcinomas, autoimmune disease, arthritis, graft rejection, inflammatory  
 CC bowel disease, proliferation induced after medical procedures, such as  
 CC surgery, angioplasty. The motor domains may also be used in  
 CC nanotechnology applications, and polynucleotides encoding the proteins  
 CC is further useful for inclusion on GeneChip (KTM) array or for use in  
 CC expression monitoring. This is the amino acid sequence of the novel human  
 CC motor protein HsKip3d motor domain, described in the method of the  
 CC invention  
 XX CC  
 XX SQ Sequence 357 AA;

Query Match 31.2%; Score 552.5; DB 5; Length 357;  
 Best Local Similarity 39.7%; Pred. No. 6.9e-48;  
 Matches 141; Conservative 59; Mismatches 114; Indels 41; Gaps 11;  
 QY 20 VRVAVRLRP-----FVDTAGASDP-----PCVRGMDSCSLEIANWRNHQ 59  
 Db 12 MKVYVVRVPENTKEKAAGFKVHVVDKHLVDPKQEEVSFFHGKTTNQNVIKQN-- 69  
 QY 60 ETLKYOFDAFYGERSTQDDIYAGSVQPIRLHLLLEGQNASVLAYGPTGAGKTHMLGSPEQ 119  
 Db 70 KDLKFVDAVDFDTSTQSEVFHTTRPILRSFLNGYNTVLTATGATGAGKTHMLGSADE 129  
 QY 120 PGVIPRALMDLLQLTRECAEGRPWALSVTMSYLEIYQEKVLDLLDPASGDLVIREDCRG 179  
 Db 130 PGVYLTMLHLKCMDEIKEE---ICSTAVSYLEVYNEQIRDLL--VNSGLAVREDTQK 185  
 QY 180 NILIPGLS-QKPISSFADFERHFLP-ASRNTVGTATRLNQRSSRSRSHAVLLVKVDQERELA 237  
 Db 186 GVVVHGLTLHQPKSS--BEILHLLDNGKNRKTQHTDMNATSSRSRSHAVFOILRQODKTA 243  
 QY 238 PFRQ--REGKLYLIDLAGSEDNRRTGNKGLRKESGAINTSLFVLGKVVVDAL-----NO 289  
 Db 244 SINQNVRIAKMSLIDLAGSERASTSGAKGTRFVEGTNINRSLALGNNVINALADSKRKQK 303  
 QY 290 GLPRVPRYRDSKLTRELQDSLGSSAHSILIANIAPERFYLDTVSALNFAARSKEV 344  
 Db 304 ---HIPYRNSKLTRELKDSLGNCQITMIAAVSPSSVFYDDTYNTLKYANRAKDI 355

Search completed: November 5, 2004, 18:44:40

Job time : 72.0157 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:57:04 ; Search time 94.4529 Seconds  
(without alignments)  
7833.865 Million cell updates/sec

Title: US-10-797-893-5  
Perfect score: 1041  
Sequence: 1 atgggtcgctgcgctaag.....ccaggaggtgatcaattga 1041

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1041	100.0	1041	3	US-09-724-224-5
2	1041	100.0	1041	4	US-10-093-317-5
3	1038.4	99.8	1464	3	US-09-724-224-7
4	1038.4	99.8	1464	4	US-10-093-317-7
5	1038	99.7	1115	3	US-09-724-224-1
6	1038	99.7	1115	4	US-10-093-317-1
7	1035.4	99.5	1538	3	US-09-724-224-3
8	1035.4	99.5	1538	4	US-10-093-317-3
9	1035.4	99.5	2097	4	US-09-595-684B-34
10	136.6	13.1	1152	4	US-09-883-096-6
11	136.6	13.1	4108	4	US-09-883-096-1
12	127.8	12.3	897	3	US-09-621-233-1
13	127.8	12.3	897	3	US-09-724-508-1
14	127.8	12.3	897	3	US-09-724-516-1
15	127.8	12.3	897	4	US-10-090-695-1
16	125.2	12.0	1026	3	US-09-641-806-1
17	125.2	12.0	1026	4	US-09-723-129-1
18	125.2	12.0	1026	4	US-09-722-862-1
19	124.2	11.9	1011	3	US-09-641-806-3
20	124.2	11.9	1011	4	US-09-723-129-3
21	124.2	11.9	1011	4	US-09-722-862-3
22	121.4	11.7	1014	4	US-09-883-096-3
23	98.6	9.5	3572	2	US-08-713-815A-2
24	93.8	9.0	3258	4	US-09-799-451-586
25	85.6	8.2	4014	3	US-09-541-782-1
26	85.6	8.2	4014	4	US-09-723-820-1
27	85.6	8.2	4014	4	US-10-270-085-1

28	79.2	7.6	1032	4	US-09-967-908A-7	Sequence 7, Appli
29	79.2	7.6	1032	4	US-10-159-151-7	Sequence 7, Appli
30	79.2	7.6	1092	4	US-09-967-908A-9	Sequence 9, Appli
31	79.2	7.6	1092	4	US-10-159-151-9	Sequence 9, Appli
32	79.2	7.6	1122	4	US-09-967-908A-3	Sequence 3, Appli
33	79.2	7.6	1122	4	US-10-159-151-3	Sequence 3, Appli
34	79.2	7.6	1146	4	US-09-967-908A-5	Sequence 5, Appli
35	79.2	7.6	1146	4	US-10-159-151-5	Sequence 5, Appli
36	79.2	7.6	6409	4	US-09-967-908A-1	Sequence 1, Appli
37	79.2	7.6	6409	4	US-10-159-151-1	Sequence 1, Appli
38	78.4	7.5	1065	3	US-09-724-511-3	Sequence 3, Appli
39	78.4	7.5	1065	4	US-09-723-097-3	Sequence 3, Appli
40	78.4	7.5	1065	4	US-09-632-344-3	Sequence 3, Appli
41	78.4	7.5	1217	3	US-09-724-511-1	Sequence 1, Appli
42	78.4	7.5	1217	4	US-09-723-097-1	Sequence 1, Appli
43	78.4	7.5	1217	4	US-09-632-344-1	Sequence 1, Appli
44	77.8	7.5	3930	3	US-09-162-373-2	Sequence 2, Appli
45	77.8	7.5	3930	3	US-09-467-946-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-724-224-5  
; Sequence 5, Application US/09724224  
; Patent No. 6387644  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/09/724, 224  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597, 292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Human  
US-09-724-224-5

Query Match	100.0%	Score 1041;	DB 3;	Length 1041;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1041;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTGCGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGCGTA	60	
Db	1	ATGGGTGCGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGCGTA	60	
QY	61	AGGGTGGCTGTGCGCTGCGGCAATTTGTGATGAAACAGCGGGAGCAAGTATCCCCC	120	
Db	61	AGGGTGGCTGTGCGCTGCGGCAATTTGTGATGAAACAGCGGGAGCAAGTATCCCCC	120	
QY	121	TGTGTGCGGGCATCGACAGCTGCTCTCTAGAGATTGCTTAAGAGAACACACAGGAG	180	
Db	121	TGTGTGCGGGCATCGACAGCTGCTCTCTAGAGATTGCTTAAGAGAACACACAGGAG	180	
QY	181	ACTCTCAATACCAAGTTTGTATGCGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT	240	
Db	181	ACTCTCAATACCAAGTTTGTATGCGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTAT	240	
QY	241	GCAGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGCAATGCCAGTGTCTT	300	
Db	241	GCAGTTTCAGTGCAGCCCATCTTAAGGCACTTGTCTGGAAGGCGCAATGCCAGTGTCTT	300	
QY	301	GCCTATGACCCACAGGAGCTGGGAGACGACACAATGCTGGGAGCCCGAGCAACT	360	
Db	301	GCCTATGACCCACAGGAGCTGGGAGACGACACAATGCTGGGAGCCCGAGCAACT	360	
QY	361	GGGTGTATCCCGCGGGCTCTCATGACCTCTGACGCTCACAAAGGAGGAGGGTCCCGAG	420	

```

Db 361 GGGGTGATCCGGGGCTCTCATGGACCTCTGAGCTCAAGAGGAGGGGTGCCGAG 420
Qy 421 GGGCGGCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAAGGTA 480
Db 421 GGGCGGCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAAGGTA 480
Qy 481 TTAGACTCTGAGACCTCTGTTGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAAT 540
Db 481 TTAGACTCTGAGACCTCTGTTGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAAT 540
Qy 541 ATCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTGTGCTATTTGAGGGGCAC 600
Db 541 ATCTGATTCGGGGTCTCTCCAGAAAGCCCATCAGTAGCTTGTGCTATTTGAGGGGCAC 600
Qy 601 TTCTGCGACCCAGTCCGAATCCGACTGTAGGAGCCACCCGGCTCAACAGAGGCTCTCTCC 660
Db 601 TTCTGCGACCCAGTCCGAATCCGACTGTAGGAGCCACCCGGCTCAACAGAGGCTCTCTCC 660
Qy 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGACCGGAGGAGGAGGAGGAGGAGGAGG 720
Db 661 CGCAGTCATGCTGTGCTCTGCTCAAGGTGGACCGGAGGAGGAGGAGGAGGAGGAGG 720
Qy 721 CAGCGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTCAAGGAGCAACCGGCGACA 780
Db 721 CAGCGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTCAAGGAGCAACCGGCGACA 780
Qy 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Qy 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Qy 901 CTCACCTGCTATTGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGGCCAAAC 960
Db 901 CTCACCTGCTATTGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGGCCAAAC 960
Qy 961 ATTGGCCCTGAGAGAGCTTCTACCTAGACAGCTCCGACTCCGACTCAACTTTGCTGCCAGG 1020
Db 961 ATTGGCCCTGAGAGAGCTTCTACCTAGACAGCTCCGACTCCGACTCAACTTTGCTGCCAGG 1020
Qy 1021 TCCAGGAGGTGATCAATTGA 1041
Db 1021 TCCAGGAGGTGATCAATTGA 1041

```

RESULT 2

```

US-10-093-317-5
; Sequence 5, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093.317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-5

```

```

Query Match 100.0%; Score 1041; DB 4; Length 1041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 ATGGGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCCGCTA 60
Db 1 ATGGGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCCGCTA 60
Qy 61 AGGGTGGCTGTGCGACTCGGCAATTTGTGATGGAACAGCGGGAGCAAGTATCCCCC 120
Db 61 AGGGTGGCTGTGCGACTCGGCAATTTGTGATGGAACAGCGGGAGCAAGTATCCCCC 120
Qy 121 TGTGTGCGGGCATGAGCAGCTGCTCTTAGAGATTTGTAATCTGAGGAAACACGAGAG 180
Db 121 TGTGTGCGGGCATGAGCAGCTGCTCTTAGAGATTTGTAATCTGAGGAAACACGAGAG 180
Qy 181 ACTCTCAAAATACCAAGTTTGTATGAGTCTTATGGGAGAGAGTACTCAGCAGCACTAT 240
Db 181 ACTCTCAAAATACCAAGTTTGTATGAGTCTTATGGGAGAGAGTACTCAGCAGCACTAT 240
Qy 241 GCAGGTTCAGTGTGAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGT 300
Db 241 GCAGGTTCAGTGTGAGCCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGT 300
Qy 301 GCCTATGAGCCACAGAGGAGCTGGGAAGACGACACAATGTGGGAGGAGGAGGAGGAG 360
Db 301 GCCTATGAGCCACAGAGGAGCTGGGAAGACGACACAATGTGGGAGGAGGAGGAGGAG 360
Qy 361 GGGGTGATTCGGCGGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGGAGGAGGAGGAG 420
Db 361 GGGGTGATTCGGCGGGCTCTCATGAGCTCTCTGAGCTCAACAGGAGGAGGAGGAGGAG 420
Qy 421 GGGCGGCATGGGCGCTTCTGTACCATGCTTCTAGAGATCTTACAGGAGTCTACAGGAG 480
Db 421 GGGCGGCATGGGCGCTTCTGTACCATGCTTCTAGAGATCTTACAGGAGTCTACAGGAG 480
Qy 481 TTAGACTCTCTGGACCTCTGCTTGGGAGAGCTGGTAATCCGAGAAGAGTCCGCGGGAAT 540
Db 481 TTAGACTCTCTGGACCTCTGCTTGGGAGAGCTGGTAATCCGAGAAGAGTCCGCGGGAAT 540
Qy 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGTAATTTGAGCGGCAC 600
Db 541 ATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGTAATTTGAGCGGCAC 600
Qy 601 TTCTGCGACCCAGTCCGAATCCGACTGTAGGAGCCACCCGGCTCAACAGGAGGAGGAG 660
Db 601 TTCTGCGACCCAGTCCGAATCCGACTGTAGGAGCCACCCGGCTCAACAGGAGGAGGAG 660
Qy 661 CGCAGTCATGCTGTGCTCTGCTGTAAGGTGGAACAGCGGGAACGTTTGGCCCAATTCGC 720
Db 661 CGCAGTCATGCTGTGCTCTGCTGTAAGGTGGAACAGCGGGAACGTTTGGCCCAATTCGC 720
Qy 721 CAGCGAGGAGGAGAACTCTACCTGATTTGACTTGGGTGGGTCAAGGAGCAACCGGCGACA 780
Db 721 CAGCGAGGAGGAGAACTCTACCTGATTTGACTTGGGTGGGTCAAGGAGCAACCGGCGACA 780
Qy 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Qy 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 841 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Qy 901 CTCACCTGCTATTGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGGCCAAAC 960
Db 901 CTCACCTGCTATTGAGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATGGCCAAAC 960
Qy 961 ATTGGCCCTGAGAGAGCTTCTACCTAGACAGCTCCGACTCCGACTCAACTTTGCTGCCAGG 1020
Db 961 ATTGGCCCTGAGAGAGCTTCTACCTAGACAGCTCCGACTCCGACTCAACTTTGCTGCCAGG 1020
Qy 1021 TCCAGGAGGTGATCAATTGA 1041
Db 1021 TCCAGGAGGTGATCAATTGA 1041

```

```

RESULT 3
US-09-724-224-7
; Sequence 7, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-7

Query Match          99.8%; Score 1038.4; DB 3; Length 1464;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTA 60
QY 61 AGGTGGCTGTGGACTCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
DB 61 AGGTGGCTGTGGACTCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
QY 121 TGTGTCGGGGATGACAGAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGG 180
DB 121 TGTGTCGGGGATGACAGAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGG 180
QY 181 ACTCTCAATATACAGTTTGTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
DB 181 ACTCTCAATATACAGTTTGTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
QY 241 GCAGGTTCAAGTGGCCCATCTTGTACCATGTCTTACCTAGAGATCTACAGGAGGTA 480
DB 241 GCAGGTTCAAGTGGCCCATCTTGTACCATGTCTTACCTAGAGATCTACAGGAGGTA 480
QY 481 TTAGACCTCTGGACCTGTTTGGGAGACCTGTTAATCCGAGAGACTGCCGGGGAT 540
DB 481 TTAGACCTCTGGACCTGTTTGGGAGACCTGTTAATCCGAGAGACTGCCGGGGAT 540
QY 541 ATCTCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTTCAGGGCAC 600
DB 541 ATCTCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTTCAGGGCAC 600
QY 601 TTCCTGCCAGCAGTGGAAATCGGACTGTAGAGGACACCCGGCTCAACAGCGCTCTCC 660
DB 601 TTCCTGCCAGCAGTGGAAATCGGACTGTAGAGGACACCCGGCTCAACAGCGCTCTCC 660
QY 661 CGCAGTCATGCTGTGCTCTGCTCAAGTGGACCGAGGAGCTTTGGCCCCATTTGCG 720
DB 661 CGCAGTCATGCTGTGCTCTGCTCAAGTGGACCGAGGAGCTTTGGCCCCATTTGCG 720

```

```

QY 721 CAGCGAGAGGGAAAACTCTACCTGATTGATTGGCTGGGTGAGAGCAACCGCGCAC 780
DB 721 CAGCGAGAGGGAAAACTCTACCTGATTGATTGGCTGGGTGAGAGCAACCGCGCAC 780
QY 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTG 840
DB 781 GGCACAAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTG 840
QY 841 GGCACAAAGTGTAGTGGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGACACG 900
DB 841 GGCACAAAGTGTAGTGGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGACACG 900
QY 901 CTCCTCGCTATTCAGGACTCTCTGGGTGGCTCAGCCACAGATATCTTATTTGCAAC 960
DB 901 CTCCTCGCTATTCAGGACTCTCTGGGTGGCTCAGCCACAGATATCTTATTTGCAAC 960
QY 961 ATTGGCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCAGG 1020
DB 961 ATTGGCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCAGG 1020
QY 1021 TCACAGGAGGTGATCAATTG 1040
DB 1021 TCACAGGAGGTGATCAATCG 1040

RESULT 4
US-10-093-317-7
; Sequence 7, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-7

Query Match          99.8%; Score 1038.4; DB 4; Length 1464;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTA 60
QY 61 AGGTGGCTGTGGACTCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
DB 61 AGGTGGCTGTGGACTCGGCCATTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120
QY 121 TGTGTCGGGGATGACAGAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGG 180
DB 121 TGTGTCGGGGATGACAGAGCTGCTCTAGAGATTGCTAACTGGAGGACCAACAGG 180
QY 181 ACTCTCAATATACAGTTTGTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
DB 181 ACTCTCAATATACAGTTTGTATGGGAGAGAGTACTCAGCAGGACATCTAT 240
QY 241 GCAGGTTCAAGTGGCCCATCTTGTACCATGTCTTACCTAGAGATCTACAGGAGGTA 300
DB 241 GCAGGTTCAAGTGGCCCATCTTGTACCATGTCTTACCTAGAGATCTACAGGAGGTA 300
QY 301 GCCTATGAGACCCACAGGAGCTGGAGAGCGCACAACTGCTGGGAGCCAGCAACT 360
DB 301 GCCTATGAGACCCACAGGAGCTGGAGAGCGCACAACTGCTGGGAGCCAGCAACT 360
QY 361 GGGGTGATCCCGGGGCTCTCATGACCTCTCTGACGCTCACAAGGAGGAGGTGCGGAG 420
DB 361 GGGGTGATCCCGGGGCTCTCATGACCTCTCTGACGCTCACAAGGAGGAGGTGCGGAG 420
QY 421 GCGCGGCCATGCGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGGTA 480
DB 421 GCGCGGCCATGCGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGGTA 480
QY 481 TTAGACCTCTGGACCTGTTTGGGAGACCTGTTAATCCGAGAGACTGCCGGGGAT 540
DB 481 TTAGACCTCTGGACCTGTTTGGGAGACCTGTTAATCCGAGAGACTGCCGGGGAT 540
QY 541 ATCTCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTTCAGGGCAC 600
DB 541 ATCTCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTTCAGGGCAC 600
QY 601 TTCCTGCCAGCAGTGGAAATCGGACTGTAGAGGACACCCGGCTCAACAGCGCTCTCC 660
DB 601 TTCCTGCCAGCAGTGGAAATCGGACTGTAGAGGACACCCGGCTCAACAGCGCTCTCC 660
QY 661 CGCAGTCATGCTGTGCTCTGCTCAAGTGGACCGAGGAGCTTTGGCCCCATTTGCG 720
DB 661 CGCAGTCATGCTGTGCTCTGCTCAAGTGGACCGAGGAGCTTTGGCCCCATTTGCG 720

```

QY 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCGACCTCCTCGACCTCACAAAGGAGAGGGTCCGAG 420  
 Db 361 GGGGTGATCCCGGGCTCTCATGGACCTCTCGACCTCCTCGACCTCACAAAGGAGAGGGTCCGAG 420  
 QY 421 GGGCGCCATGGGCGCCCTTCTGTACCATGTCTTACCTAGAGATCTACCGAGAGAGGTA 480  
 Db 421 GGGCGCCATGGGCGCCCTTCTGTACCATGTCTTACCTAGAGATCTACCGAGAGAGGTA 480  
 QY 481 TTAGACCTCTGGACCTCTGCTTGGGAGACCTGGTAATCCGAGAGAGCTGCGGGGAAT 540  
 Db 481 TTAGACCTCTGGACCTCTGCTTGGGAGACCTGGTAATCCGAGAGAGCTGCGGGGAAT 540  
 QY 541 ATCTGATTTCCGGGCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGCGGCAC 600  
 Db 541 ATCTGATTTCCGGGCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGCGGCAC 600  
 QY 601 TTCTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACCGAGGCTCTCC 660  
 Db 601 TTCTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACCGAGGCTCTCC 660  
 QY 661 CGCAGTCATGCTGTCTGTCTGTAAGGTGGACCGGGAACGTTTGGCCCCCATTTGCG 720  
 Db 661 CGCAGTCATGCTGTCTGTCTGTAAGGTGGACCGGGAACGTTTGGCCCCCATTTGCG 720  
 QY 721 CAGCGAGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGGCGCAC 780  
 Db 721 CAGCGAGAGGGAATCTTACCTGATTTGACTTGGCTGGGTGAGAGCAACCGGCGCAC 780  
 QY 781 GSCAAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTG 840  
 Db 781 GSCAAAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTTGTCTG 840  
 QY 841 GSCAAAGTGTAGTGGCTGTAATCAGGGCTCCTCGTGTACCTTATCGGGAAGCAAG 900  
 Db 841 GSCAAAGTGTAGTGGCTGTAATCAGGGCTCCTCGTGTACCTTATCGGGAAGCAAG 900  
 QY 901 CTCACCTCGCTATTCAGAGCTCTCTGGTGGCTCAGCCCAAGTATCTTATGSCAAC 960  
 Db 901 CTCACCTCGCTATTCAGAGCTCTCTGGTGGCTCAGCCCAAGTATCTTATGSCAAC 960  
 QY 961 ATTGCCCTGAGAGAGCTTCTACCTAGACAGCTCTCGGCTCAGCTTCTGCTGCCAGG 1020  
 Db 961 ATTGCCCTGAGAGAGCTTCTACCTAGACAGCTCTCGGCTCAGCTTCTGCTGCCAGG 1020  
 QY 1021 TCCAGGAGGTGATCAATTG 1040  
 Db 1021 TCCAGGAGGTGATCAATCG 1040

RESULT 5

US-09-724-224-1  
 ; Sequence 1, Application US/09724224  
 ; Patent No. 6387644  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: No. 6387644el motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/09/724,224  
 ; CURRENT FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1115  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-724-224-1

Query Match 99.7%; Score 1038; DB 3; Length 1115;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCTCACTCCAGCTCGGTAAAGG 63  
 Db 78 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCTCACTCCAGCTCGGTAAAGG 137  
 QY 64 GTGGCTGTGGCACTGGCGCCATTTGTGGATGGAAACGGGGAGCAAGTGATCCCGCTGT 123  
 Db 138 GTGGCTGTGGCACTGGCGCCATTTGTGGATGGAAACGGGGAGCAAGTGATCCCGCTGT 197  
 QY 124 GTGGGGGCATGGACAGCTGCTCTCTAGAGATTGTCTAACTGGAGAAACCAACAGAGACT 183  
 Db 198 GTGGGGGCATGGACAGCTGCTCTCTAGAGATTGTCTAACTGGAGAAACCAACAGAGACT 257  
 QY 184 CTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCTACAGGACATCTATGCA 243  
 Db 258 CTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCTACAGGACATCTATGCA 317  
 QY 244 GGTTCAGTGCAGCCATCCTTAAGCAGCTTGTCTGGAAGGCGAGATGCCAGTGTCTTGGC 303  
 Db 318 GGTTCAGTGCAGCCATCCTTAAGCAGCTTGTCTGGAAGGCGAGATGCCAGTGTCTTGGC 377  
 QY 304 TATGGAACCCACAGAGAGCTGGGAAGACGACACAATGTCTGGGAGGAGGAGGAGGAGG 363  
 Db 378 TATGGAACCCACAGAGAGCTGGGAAGACGACACAATGTCTGGGAGGAGGAGGAGGAGG 437  
 QY 364 GTGATCCCGCGGGCTCTCATGGAACCTCTGTCAGCTCAACAGGAGAGGAGGAGGAGG 423  
 Db 438 GTGATCCCGCGGGCTCTCATGGAACCTCTGTCAGCTCAACAGGAGAGGAGGAGGAGG 497  
 QY 424 CGGCCATGGGCGCTTCTGTCAACCATGTCTTACCTAGAGATCTACAGGAGAGGATTA 483  
 Db 498 CGGCCATGGGCGCTTCTGTCAACCATGTCTTACCTAGAGATCTACAGGAGAGGATTA 557  
 QY 484 GACCTCTCGGACCTGTCTCGGAGAGCTGTGTAACTCGAGAGAGCTGCGGGGAGATATC 543  
 Db 558 GACCTCTCGGACCTGTCTCGGAGAGCTGTGTAACTCGAGAGAGCTGCGGGGAGATATC 617  
 QY 544 CTGATTCGGGCTCTCTCCAGAGCCCATCATGATGCTTGTCTGATTTTGAGCGGCACTTC 603  
 Db 618 CTGATTCGGGCTCTCTCCAGAGCCCATCATGATGCTTGTCTGATTTTGAGCGGCACTTC 677  
 QY 604 CTGCGAGCCAGTGGAAATCGGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 663  
 Db 678 CTGCGAGCCAGTGGAAATCGGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC 737  
 QY 664 AGTCATGCTGTCTCTGGTCAAGGTGGAACAGCGGGAACGTTTGGGCCCAATTTGCCAG 723  
 Db 738 AGTCATGCTGTCTCTGGTCAAGGTGGAACGTTTGGGCCCAATTTGCCAG 797  
 QY 724 CGAGAGGGAAACTCTTACCTGATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGG 783  
 Db 798 CGAGAGGGAAACTCTTACCTGATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGG 857  
 QY 784 AACAGGGCCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCTGTTTGTCTTGGC 843  
 Db 858 AACAGGGCCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCTGTTTGTCTTGGC 917  
 QY 844 AAAGTGTAGATGCGCTGGAATCAGGGCTCCTCTGCTGTATCTTATCGGGAGAGCAAGTC 903  
 Db 918 AAAGTGTAGATGCGCTGGAATCAGGGCTCCTCTGCTGTATCTTATCGGGAGAGCAAGTC 977  
 QY 904 ACTCGCCATTTGAGGAGCTCTCTGGGTGGCTCAGCCCAACAGTATCTTATGCGCAATTT 963  
 Db 978 ACTCGCCATTTGAGGAGCTCTCTGGGTGGCTCAGCCCAACAGTATCTTATGCGCAATTT 1037  
 QY 964 GCCCTGAGAGAGCTTCTACCTAGACAGAGTCTCCGCACTCAACTTTTGTGCGAGTCC 1023  
 Db 1038 GCCCTGAGAGAGCTTCTACCTAGACAGAGTCTCCGCACTCAACTTTTGTGCGAGTCC 1097  
 QY 1024 AAGGAGGTGATCAATTGA 1041  
 Db 1098 AAGGAGGTGATCAATTGA 1115

RESULT 6					
US-10-093-317-1					
; Sequence 1, Application US/10093317					
; Patent No. 6762043					
; GENERAL INFORMATION:					
; APPLICANT: Beraud, Christophe					
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for					
; FILE REFERENCE: 1044					
; CURRENT APPLICATION NUMBER: US/10/093,317					
; PRIOR FILING DATE: 2002-03-06					
; PRIOR APPLICATION NUMBER: 09/724,224					
; PRIOR FILING DATE: 2000-11-28					
; NUMBER OF SEQ ID NOS: 8					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 1					
; LENGTH: 1115					
; TYPE: DNA					
; ORGANISM: Human					
US-10-093-317-1					
Query Match					
Best Local Similarity 99.7%; Score 1038; DB 4; Length 1115;					
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	4	GTCGCTGTGGCTAAGCAAGATTGGAGTACTCGTGCCACCTCCAGCTCCGGTAAGG	63		
Dd	78	GTCGCTGTGGCTAAGCAAGATTGGAGTACTCGTGCCACCTCCAGCTCCGGTAAGG	137		
Qy	64	GTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	123		
Dd	138	GTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	197		
Qy	124	GTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	183		
Dd	198	GTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	257		
Qy	184	CTCAATACCAGTTTGCATGCCCTTCTATGGGAGAGAGTACTCAGCAGCATCTATGCA	243		
Dd	258	CTCAATACCAGTTTGCATGCCCTTCTATGGGAGAGAGTACTCAGCAGCATCTATGCA	317		
Qy	244	GGTTCAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	303		
Dd	318	GGTTCAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT	377		
Qy	304	TATGGACCCACAGGAGCTGGAGAGCGCACAAATGCTGGCAGCCACAGACCTGGG	363		
Dd	378	TATGGACCCACAGGAGCTGGAGAGCGCACAAATGCTGGCAGCCACAGACCTGGG	437		
Qy	364	GTGATCCCGGGCTCTCATGACCTCTCTGACCTCTGACCTCACAAGGGAGGGTCCGAGGGC	423		
Dd	438	GTGATCCCGGGCTCTCATGACCTCTCTGACCTCTCAGAGCTCACAAGGGAGGGTCCGAGGGC	497		
Qy	424	CGGCCATGGGCCCTTCTGTCAACATGTTTACCTAGAGATCTACAGGAGAAGTATTA	483		
Dd	498	CGGCCATGGGCCCTTCTGTCAACATGTTTACCTAGAGATCTACAGGAGAAGTATTA	557		
Qy	484	GACCTCTGGACCTCTCTGGGAGACCTGGTAAATCCGAGAGACTGCGGGGAATATC	543		
Dd	558	GACCTCTGGACCTCTCTGGGAGACCTGGTAAATCCGAGAGACTGCGGGGAATATC	617		
Qy	544	CTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGGCGGCATTC	603		
Dd	618	CTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGGCGGCATTC	677		
Qy	604	CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	663		
Dd	678	CTGCCAGCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	737		
Qy	664	AGTCATGCTGTGCTCTGTGTAAGGTGGACCAAGCGGAAAGTGTGGCCCCATTTGGCCAG	723		

```

Db 378 TATGACCCACAGAGAGCTGGAGAGGACACACATGCTGGGAGCCGACAGAGCAACTGGG 437
QY 364 GTGATCCCGCGGCTCTCATGACCTCTGAGCTTCAAGAGGAGAGGAGGAGGAGGAGGAGG 423
Db 438 GTGATCCCGCGGCTCTCATGACCTCTGAGCTTCAAGAGGAGAGGAGGAGGAGGAGGAGG 497
QY 424 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGATATTA 483
Db 498 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGATATTA 557
QY 484 GACCTCTCGGACCTGCTTCGGGAGACCTGTGTAATCCGAGAGAGCTGCCGGGGAATATC 543
Db 558 GACCTCTCGGACCTGCTTCGGGAGACCTGTGTAATCCGAGAGAGCTGCCGGGGAATATC 617
QY 544 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGCACTTC 603
Db 618 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGGCACTTC 677
QY 604 CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGGAGGCTCTCCGCG 663
Db 678 CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGGAGGCTCTCCGCG 737
QY 664 AGTCATGCTGCTCTGCTCAAGGTGGACCGGAGAACGTTTGGCCGCCATTTCCGCCAG 723
Db 738 AGTCATGCTGCTCTGCTCAAGGTGGACCGGAGAACGTTTGGCCGCCATTTCCGCCAG 797
QY 724 CGAGAGGGAATACTTACCTGATGACTTGGTGGGTGAGAGCAACCGGCGCACAGGC 783
Db 798 CGAGAGGGAATACTTACCTGATGACTTGGTGGGTGAGAGCAACCGGCGCACAGGC 857
QY 784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTGCTCTGGGC 843
Db 858 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTGTTCTCTGGGC 917
QY 844 AAAGTGTAGTGGCTGAATCAGAGGCTTCCCTGCTGATCTTATCGGAGCAGCAAGCTC 903
Db 918 AAAGTGTAGTGGCTGAATCAGAGGCTTCCCTGCTGATCTTATCGGAGCAGCAAGCTC 977
QY 904 ACTCGCTTATTCAGAGCTCTCTGGGAGCTCAGCCCAAGATCTTATTCGCAACATT 963
Db 978 ACTCGCTTATTCAGAGCTCTCTGGGAGCTCAGCCCAAGATCTTATTCGCAACATT 1037
QY 964 GCCCTGAGAGAGCTTCTACCTAGACAGCTCCGCACTCAACTTTGCTGCCAGTCC 1023
Db 1038 GCCCTGAGAGAGCTTCTACCTAGACAGCTCCGCACTCAACTTTGCTGCCAGTCC 1097
QY 1024 AAGGAGGTGATCAATTG 1040
Db 1098 AAGGAGGTGATCAATCG 1114

```

RESULT 8

```

US-10-093-317-3
; Sequence 3, Application US/10093317
; Patent No. 6762043
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10/093,317
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/724,224
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Human
US-10-093-317-3

```

```

Query Match 99.5%; Score 1035.4; DB 4; Length 1538;
Best Local Similarity 99.9%; Pred. No. 9.9e-315;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTCTCGGCTAAAGCAAGATTGGAGCTACTCTGCTGCTCCACCTCCAGCTCGCTAAGG 63
Db 78 GGTGCTCTCGGCTAAAGCAAGATTGGAGCTACTCTGCTGCTCCACCTCCAGCTCGCTAAGG 137
QY 64 GTGGCTGTGCACTGCGGCCCATTTGTGATGGAACAGCGGAGCAAGTGTATCCCTGCTGT 123
Db 138 GTGGCTGTGCACTGCGGCCCATTTGTGATGGAACAGCGGAGCAAGTGTATCCCTGCTGT 197
QY 124 GTGGGCGCATGGAGCAGCTCTCTAGAGATTGCTAACTCGAGAAACACAGAGAGACT 183
Db 198 GTGGGCGCATGGAGCAGCTCTCTAGAGATTGCTAACTCGAGAAACACAGAGAGACT 257
QY 184 CTCAATAACCAAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243
Db 258 CTCAATAACCAAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317
QY 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGTAAGGGCAGATGCGCAGTGTGCTGCC 303
Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTGTGTAAGGGCAGATGCGCAGTGTGCTGCC 377
QY 304 TATGACCCACAGAGAGCTGGAGACGACACAATGCTGGGAGAGGAGGAGGAGGAGGAGG 363
Db 378 TATGACCCACAGAGAGCTGGAGAGCAGCACAATGCTGGGAGAGGAGGAGGAGGAGGAGG 437
QY 364 GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTTCAAGGGAGAGGAGGAGGAGGAGG 423
Db 438 GTGATCCCGCGGCTCTCATGGAACCTCTGAGCTTCAAGGGAGAGGAGGAGGAGGAGG 497
QY 424 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483
Db 498 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 557
QY 484 GACCTCTCGGAGCTTCTGCTGGAGAGCTGTAATCGAGAGAGCTCCGCGGGAATATC 543
Db 558 GACCTCTCGGAGCTTCTGCTGGAGAGCTGTAATCGAGAGAGCTCCGCGGGAATATC 617
QY 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGGAGCGGCACTTC 603
Db 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGGAGCGGCACTTC 677
QY 604 CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGGAGCTCTCCGCG 663
Db 678 CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCGGCTCAACAGGAGCTCTCCGCG 737
QY 664 AGTCATGCTGCTCTGCTCAAGGTGGACCGGAGAACGTTTGGGCGGCACTTTCCGCG 723
Db 738 AGTCATGCTGCTCTGCTCAAGGTGGACCGGAGAACGTTTGGGCGGCACTTTCCGCG 797
QY 724 CGAGAGGGAATACTTACCTGATGACTTGGTGGGTGAGAGCAACCGGCGCACAGGC 783
Db 798 CGAGAGGGAATACTTACCTGATGACTTGGTGGGTGAGAGCAACCGGCGCACAGGC 857
QY 784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTGCTCTGGGC 843
Db 858 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCTGTTGTTCTCTGGGC 917
QY 844 AAAGTGTAGATGCGCTGAATCAGGGCTCCTCTGCTGATCTTATCGGAGCAGCAAGCTC 903
Db 918 AAAGTGTAGATGCGCTGAATCAGGGCTCCTCTGCTGATCTTATCGGAGCAGCAAGCTC 977
QY 904 ACTCGCTTATTCAGAGCTCTCTGGGAGCTCAGCCCAAGATCTTATTCGCAACATT 963
Db 978 ACTCGCTTATTCAGAGCTCTCTGGGAGCTCAGCCCAAGATCTTATTCGCAACATT 1037
QY 964 GCCCTGAGAGAGCTTCTACCTAGACAGCTCCGCACTCAACTTTGCTGCCAGTCC 1023
Db 1038 GCCCTGAGAGAGCTTCTACCTAGACAGCTCCGCACTCAACTTTGCTGCCAGTCC 1097
QY 1024 AAGGAGGTGATCAATTG 1040

```



Db 1098 AAGAGGTGATCAATCG 1114  
 |||||  
 RESULT 9  
 US-09-595-684B-34  
 ; Sequence 34, Application US/09595684B  
 ; Patent No. 6544766  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Ohashi, Caxa  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Vaisberg, Eugeni  
 ; APPLICANT: Wood, Kenneth  
 ; APPLICANT: Yu, Ming  
 ; TITLE OF INVENTION: Human kinesins and methods of producing  
 ; TITLE OF INVENTION: and purifying human kinesins  
 ; FILE REFERENCE: cytop036  
 ; CURRENT APPLICATION NUMBER: US/09/595,684B  
 ; PRIOR FILING DATE: 2002-06-24  
 ; PRIOR APPLICATION NUMBER: 09/295,612  
 ; PRIOR FILING DATE: 2000-04-20  
 ; NUMBER OF SEQ ID NOS: 105  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 34  
 ; LENGTH: 2097  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-595-684B-34  
 Query Match 99.5%; Score 1035.4; DB 4; Length 2097;  
 Best Local Similarity 99.9%; Pred. No. 9.9e-315;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GTGCGCTGTCGCTAGCAAGATTGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 63  
 DB 100 GTGCGCTGTCGCTAGCAAGATTGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 159  
 QY 64 GTGCGCTGTCGCTAGCAAGATTGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 123  
 DB 160 GTGCGCTGTCGCTAGCAAGATTGAGCTACTGCTGCTCCACCTCCAGCTCGCGTAAGG 219  
 QY 124 GTGCGCGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACCCACAGGAGACT 183  
 DB 220 GTGCGCGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACCCACAGGAGACT 279  
 QY 184 CTCMAATACAGTTGATGCTCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243  
 DB 280 CTCMAATACAGTTGATGCTCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 339  
 QY 244 GGTTCAGTCAGCCCATCTAAGGCACTTGTGGAAGGCGAGAATGCCAGTGTCTTGCC 303  
 DB 340 GGTTCAGTCAGCCCATCTAAGGCACTTGTGGAAGGCGAGAATGCCAGTGTCTTGCC 399  
 QY 304 TATGGAACCAAGAGCTGGGAAGACGACCAATGCTGGGACGCCAGAGCAACCTGGG 363  
 DB 400 TATGGAACCAAGAGCTGGGAAGACGACCAATGCTGGGACGCCAGAGCAACCTGGG 459  
 QY 364 GTGATCCCGGGGCTCTATGACCTCTGAGCTCACAAGGGAGGGGTGCCAGGGC 423  
 DB 460 GTGATCCCGGGGCTCTATGACCTCTGAGCTCACAAGGGAGGGGTGCCAGGGC 519  
 QY 424 CGGCATCGGGCTTCTGTCACCATGCTTTACCTAGAGATCTACAGGAGAGGTATTA 483  
 DB 520 CGGCATCGGGCTTCTGTCACCATGCTTTACCTAGAGATCTACAGGAGAGGTATTA 579  
 QY 484 GACCTCTGGAACCTGCTTCGGGAGACTGGTAATCCGAGAAGACTCCCGGGGGAATATC 543  
 DB 580 GACCTCTGGAACCTGCTTCGGGAGACTGGTAATCCGAGAAGACTCCCGGGGGAATATC 639  
 QY 544 CTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGAAGCGCACTTC 603  
 DB 640 CTGATTCGGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGAAGCGCACTTC 699

```

Db 274 AGGCTCTGGACAGCTTCTCTCAGGGCTACACTGCTCAGTGTGTGCTACGGGGCCACC 333
Qy 316 GGAGCTGGAGAGACGACACATGCTGGGAGCCAGAGAACCTGGGGGTGATCCCGGG 375
Db 334 GGGCTGGGAAGACACACACCTGCTGGGAAGGGGGGACCCCGGCATCATGTA---- 389
Qy 376 GCTCTCATGACCTCTGAGCTCACAAGGAGGAGGTGCCGAGGGCCGGCCATGGGCC 435
Db 390 -----CCTGACCACCGTGGAACTGTACAGGCGCTGGAGGCCCGCCAGCAGAGAGAAC 444
Qy 436 CTCTCTGTCACCATGCTTACCTTAGAGATCTACAGGAGAAAGTATTAGACCTCTCTGGAC 495
Db 445 TTGAGGTGCTCATCAGCTACAGGAGGTGTATATGACACATCCATGACCTCTCTGGAG 504
Qy 496 CTGCTTTCGGGAGACCTGTGTAATCCGAGAGACTGCGGGGGAAATCTCTGATTCGGGT 555
Db 505 C---CCAAGGGGCCCCCTTGCATCCGAGAGACCCCGACAAGGGGGTGGTGAAGA 561
Qy 556 CTCTCCAGAGCCCATCAGTGTGCTGATTTTGAGGGGCACTTCTCTGCCAGCAGT 615
Db 562 CTCTCTTTCCACGAGCCAGCTCAGCCGAGAGCTGCTGGAGATATGACAGGGGGAAC 621
Qy 616 CGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCGAGTCAATGCTGTG 675
Db 622 CGTAACCGCAGCAGACACCCCACTGATGCCAAGCGACTTCTCTCGCTCCCATGCCATC 681
Qy 676 CTCTGCTCAAGGTGACAGCGGGGAACTGTTGGCCCCAATTCGCGAG-----CGAGAG 729
Db 682 TTCCAGATCTTTGTGAAGCAGCAGGACCGGGTTCCAGGACTGACCCAGGCTGTCCAGGTG 741
Qy 730 GGAATACTTACTGATTTGCTGCTGCTGAGGACACCGGGCAGAGCAACAG 789
Db 742 GCAAGATGAGCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
Qy 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTGCTCTGGGCAAGTG 849
Db 802 GGGAGCGCTGCGGAGGGGCGCAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 861
Qy 850 GTAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 862 CTCAATGCTTGGCGGATCAAGAGGGCGCAAGACCCATGTGCTGCTGCTGCTGCTGCTGCTGCTG 921
Qy 901 CTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 922 CTGACCGGCTGCTCAAGACTTCTCTGGGGGCACTGCGGCGACAGTATGATCGTCC 981
Qy 961 ATTGCCCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
Db 982 ATCAGCCCCCTCAGCGCTGACCTACGAGGACAGTACACACCTTCAATATCCCGACCGG 1041
Qy 1021 TCAAGGAGGTGA 1033
Db 1042 GCCAAGGAGATCA 1054

```

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid sequence of human kinesin motor
; OTHER INFORMATION: protein gene Hskip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a gene.
US-09-883-096-1

Query Match 13.1%; Score 136.6; DB 4; Length 4108;
Best Local Similarity 51.2%; Pred. No. 2.4e-32; Indels. 27; Gaps 4;
Matches 437; Conservative 0; Mismatches 389;

Qy 136 TTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCAAGTTTCAGTGGC 255
Db 344 TTTGACCGGTCTTTTGGCGAGCGGCCACCAACAGGACGTTTCCAGCACACCGCAC 403
Qy 256 CCATCTCTAAGGCACTTCTGTAAGGGGAGAAATGCGAGTGTGCTTGTCTATGAGCCACA 315
Db 404 AGCGTCTCTGAGCAGCTTCTCCAGGGCTACAACCTGCTCAGTGTGTTTGCCTACGGGGCCACC 463
Qy 316 GGAGCTGGAGAGACGACACACAATGCTGGGAGCCAGAGCAACCTGGGGTGTATCCCGGG 375
Db 464 GGGGCTGGAGAGACACACCAATGCTGGAGAGGGAGGGGACCCCGGATCATGTA---- 519
Qy 376 GCTCTCATGAGACCTCTCTGAGCTCAAGGGAGGAGGTGCGGAGGGCGGCGCATGGGCC 435
Db 520 -----CCTGACCACCGTGGAACTGTACAGGCGCTGGAGGCCCGCCAGCAGGAGAAC 574
Qy 436 CTCTCTGCTGACATGCTTACCTAGAGATCTACAGGAGAGAGTATTAGACCTCTCTGAGC 495
Db 575 TTCGAGGTGCTCATCAGCTACAGGAGGTGTATATGAACAGATCATGACCTCTCTGGAG 634
Qy 496 CTGCTCTCGGAGACCTGCTAATCCGAGAGAGTCCGCGGGGAAATCTCTGATTCGGGT 555
Db 635 C---CCAAGGGGCCCCCTTGCATCCGCGAGACCCCGACAGGGGGTGGTGTGAAGA 691
Qy 556 CTCTCCAGAGCCCATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
Db 692 CTCTCTTCTCCAGCAGCCAGCTCAGCGAGCAGCTGCTGGAGATACTGACAGGGGGAAC 751
Qy 616 CGAATCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGAGTATGCTGTG 675
Db 752 CGTAACCGCAGCAGCAGCAGCCCACTGATGCCAACCGGACTTCTCTCCGCTCCCATGCCATC 811
Qy 676 CTCTCTGCTCAAGGTGAGCAGCGGAGACGTTTGGGCCCATTTGCGCAG-----CGAGAG 729
Db 812 TTCCAGATCTTTGTAAGCAGAGGACCGGGTTCAGGAGTACCCAGGCTGTCCAGGTG 871
Qy 730 GGAATACTCTA CCTGATGACTTGTGCTGGGTGAGAGCAACCGGCGCAGAGGCAACAG 789
Db 872 GCCAAGATGAGCTGATTGACCTGGCTCAGAGCGGGCATCCAGCACCCATCGCAAG 931
Qy 790 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTCTTGTGCTGGGCAAGTG 849
Db 932 GGGGAGCGGCTGCGGGAGGGGCCAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 991
Qy 850 GTAGATGCGCTGAATCAGG-----GCCTCCCTCGTGTACCTTATCGGAGACAGCAAG 900
Db 992 CTCAATGCTTTGGCGGATGCAAGGGCCGCAAGACCCATGTGCCCTACCGGAGACGAAA 1051
Qy 901 CTCACCTGCTTATGAGAGTCTCTGCGGTGGCTCAGCCCAAGTATCTTATTGCCAAC 960
Db 1052 CTGACCGGCTGCTCAAGAGACTCTCTCCGGGGCAACTGCGGCGACAGTATGATCGCTGCC 1111
Qy 961 ATTGCCCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020
Db 1112 ATCAGCCCCCTCAGCGCTGACCTACGAGGACAGTACAAACCCCTCAATATATCCCGACCGG 1171
Qy 1021 TCCAAGGAGGTGA 1033

```

```

RESULT 11
US-09-883-096-1
; Sequence 1, Application US/09883096
; Patent No. 6680369
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; PRIORITY FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15

```

```
Db 1172 GCCAAGGAGATCA 1184
|||||
RESULT 12
US-09-621-233-1
; Sequence 1, Application US/09621233
; Patent No. 6294371
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6294371el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/09/621.233
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (472)...(495)
; OTHER INFORMATION: n = a, c, t, or g
US-09-621-233-1

Query Match 12.3%; Score 127.8; DB 3; Length 897;
Best Local Similarity 50.1%; Pred. No. 5.9e-30;
Matches 429; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 190 TACCAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGGTTCA 249
Db 106 TACCTGTTCAGTGGCTTTGACTTACCGCCACCCAGAGATGGTGTATCAGGCCACC 165
Qy 250 GTGCAGGCCATCTTAAGGCACCTTGTGGAAGGAGGAGTACTCAGCAGGACATCTATGAGGTTCA 249
Db 166 ACCAAGAGCTCATCAGGGCGTCTATCAGGCTACATGAGGATGATCCGATGGC 225
Qy 310 CCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCAGCAGCACTTGGGTGATC 369
Db 226 CCCACAGGCTGTGGGAAACCTACACCATCTGCGGACACAGACCAGGAGCTGATCAT 285
Qy 370 CCGCGGCTCTCATGAGCTTCCCTGAGCTCACAAGGAGGAGGAGTGCAGGAGCCGCCA 429
Db 286 GTTCAGACCTTCAACGACCTTCTCCGTGCATCGAGGAGACCAATGA-----C 336
Qy 430 TGGGCCCTTTCTGTCAACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTC 489
Db 337 ATGAGTATGAGTCTCCATGCTCTACTCTGGAGATCTACAAATGAGATGATCCGGACCTG 396
Qy 490 CTGGAACCTTCTTCCGGAGACCTTGGTAATCCGAGAGAGACTGCCGGGGGAATATCTGATT 549
Db 397 CTGAACCCCTCCCTGGGCTACCTGGAGCTCGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456
Qy 550 CCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTCCCTGCCA 609
Db 457 GCGGCAATACCGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Qy 610 GCCAGTCAAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGCGCTCTCCCGCAGTCAAT 669
Db 517 GGGAAACGGCAGAGGACCCAGGAGCCACCGCCGACCAACAGAGCTCTCCGCTCCAC 576
Qy 670 GCTGTGCTCTCTGATAGGTGAGACCGCGGGAAGCTTTGGCCCAATTTCCAGAGGAGAG 729
Db 577 GCGGGCGCGCTGTCA----- 592
Qy 730 GGAATACTCTACCTGATTGCTGGCTGGGTGAGGAGCAACCGCGCAGCAGCAACAG 789
Db 593 -----TGATCGACCTTGGCTTGGCTCAGAGCGGCCCTCGCAGACACAGAAATCGT 639
Qy 790 GGCCTTTGGCTAAAGAGAGTGGAGCCATCAACACCTTCCCTGTTTGTCTGGGCAAGTG 849
|||||
Db 640 GGCAGCGTATGAAGGAGGGGGGCCACATCAACCGCTCACTGCTGGCACTGGGCAATGC 699
Qy 850 GTAGATGCGCTGAATCAG-----GGCCTCCTGCTGTACCTTATCGGAGCAGCAAGCTC 903
Db 700 ATCAACGCCCTGAGCGACAAGGGTAGCAACAGTACATCACTATCCGACAGCAAGCTC 759
Qy 904 ACTCGCCTATTGAGGAGCTCTCTGGGTGGCTCAGCCCAACAGTATCTCTTATTGCCAACATT 963
Db 760 ACCGGCTCTCTGAAGGACTCTCTGGGAGGAAACAGCCGACAGTATGATCGCTCACATC 819
Qy 964 GCCCCTGAGAGCGCTTCTTACCTTAGACACAGTCTCCGACATCACTTTGTCGCAAGTCC 1023
Db 820 AGTCTCGAGCAGTGCCTTCGAGGAGTCCCGAAACACCTGACCTACGCGCGCGGCC 879
Qy 1024 AAGGAGTGATCAATTG 1040
Db 880 AAGAACATTAAAGACTAG 896
|||||

RESULT 13
US-09-724-508-1
; Sequence 1, Application US/09724508
; Patent No. 6368841
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6368841el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/09/724.508
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/621.233
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (472)...(495)
; OTHER INFORMATION: n = a, c, t, or g
US-09-724-508-1

Query Match 12.3%; Score 127.8; DB 3; Length 897;
Best Local Similarity 50.1%; Pred. No. 5.9e-30;
Matches 429; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 190 TACCAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGGTTCA 249
Db 106 TACCTGTTCAGTGGCTTTGACTTACCGCCACCCAGAGATGGTGTATCAGGCCACC 165
Qy 250 GTGCAGGCCATCTTAAGGCACCTTGTGGAAGGAGGAGTACTCAGCAGGACATCTATGAGGTTCA 249
Db 166 ACCAAGAGCTCATCAGGGCGTCTATCAGGCTACATGAGGATGATCCGATGGC 225
Qy 310 CCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCCAGCAGCACTTGGGTGATC 369
Db 226 CCCACAGGCTGTGGGAAACCTACACCATCTGCGGACACAGACCAGGAGCTGATCAT 285
Qy 370 CCGCGGCTCTCATGAGCTTCCCTGAGCTCACAAGGAGGAGGAGTGCAGGAGCCGCCA 429
Db 286 GTTCAGACCTTCAACGACCTTCTCCGTGCATCGAGGAGACCAATGA-----C 336
Qy 430 TGGGCCCTTTCTGTCAACCATGTCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTC 489
Db 337 ATGAGTATGAGTCTCCATGCTCTACTCTGGAGATCTACAAATGAGATGATCCGGACCTG 396
Qy 490 CTGGAACCTTCTTCCGGAGACCTTGGTAATCCGAGAGAGACTGCCGGGGGAATATCTGATT 549
Db 397 CTGAACCCCTCCCTGGGCTACCTGGAGCTCGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456
Qy 550 CCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGTAGCGGCACTTCCCTGCCA 609
Db 457 GCGGCAATACCGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Qy 610 GCCAGTCAAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGCGCTCTCCCGCAGTCAAT 669
Db 517 GGGAAACGGCAGAGGACCCAGGAGCCACCGCCGACCAACAGAGCTCTCCGCTCCAC 576
Qy 670 GCTGTGCTCTCTGATAGGTGAGACCGCGGGAAGCTTTGGCCCAATTTCCAGAGGAGAG 729
Db 577 GCGGGCGCGCTGTCA----- 592
Qy 730 GGAATACTCTACCTGATTGCTGGCTGGGTGAGGAGCAACCGCGCAGCAGCAACAG 789
Db 593 -----TGATCGACCTTGGCTTGGCTCAGAGCGGCCCTCGCAGACACAGAAATCGT 639
Qy 790 GGCCTTTGGCTAAAGAGAGTGGAGCCATCAACACCTTCCCTGTTTGTCTGGGCAAGTG 849
|||||
```



; LENGTH: 897  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (472)..  
; OTHER INFORMATION: n = a, c, t, or g  
US-10-090-695-1

Search completed: November 10, 2004, 18:26:02  
Job time : 99.4529 secs

```
Query Match      12.3%; Score 127.8; DB 4; Length 897;
Best Local Similarity 50.1%; Pred. No. 5.9e-30;
Matches 429; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 190 TACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCAGGTTCA 249
Db      |||||
Qy 106 TACCTGTTGACGTGGCTTTGACTTCACCGCCACCAGGAGATGGTGTATCAGGCCACC 165
Db      |||||
Qy 250 GTGCAGCCCATCTTAAGCACTTCTCGAAGGGCAGATGCCAGTGTGCTTGCTATGGA 309
Db      |||||
Qy 166 ACCAAGAGCTCATCGAGGGCGTCTCTCAGGCTACAATGCCACTGCTTTGGCTATGCC 225
Db      |||||
Qy 310 CCCACAGGAGCTGGGAAGAGCACAACAATCTGGGCGAGCCACAGAGCAACCTGGGGTGATC 369
Db      |||||
Qy 226 CCCACAGGCTGTGGGAAAACTACACCATGCTGGGCAACAGCAGGAGCCTGGCATCTAT 285
Db      |||||
Qy 370 CCAGGGCTCTCATGGACCTCTCTGACCTCACAAAGGAGGAGGGTCCGAGGGCCGCCCA 429
Db      |||||
Qy 286 GTTCAGACCTCTCAACGACCTCTTCGGTCCATCGAGGAGACCAGCAATGA-----C 336
Db      |||||
Qy 430 TGGGCCCCCTTCTGTACCATGTCTTACCTAGAGATCTACAGGAGAGGATTTAGACCTC 489
Db      |||||
Qy 337 ATGGAGTATGAGGTCTCCATGTCTACCTGGAGATCTACAATGAGATGATCCGGGACCTG 396
Db      |||||
Qy 490 CTGACCCCTCTCGGAGACCTGTTAATCCGAGAAGCTGCCGGGGGAATATCCTGATT 549
Db      |||||
Qy 397 CTGNAACCCCTCCCTGGGCTACTGGAGCTGGGGAGGACTTAAGGGGGTGTCCAGGTG 456
Db      |||||
Qy 550 CCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTGAGCGGCACCTTCCTGCCA 609
Db      |||||
Qy 457 GCCGGCATCACGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Db      |||||
Qy 610 GCCAGTCGAATCGGACTGTAGAGGCCACCGGGTCAACAGAGCGCTCTCCCGAGTCAAT 669
Db      |||||
Qy 517 GGGAAACCGGAGAGGACCCAGGAGCCCAACCGCCGCCAACAGAGAGCTCTCCCGCTCCAC 576
Db      |||||
Qy 670 GCTGTGCTCTCTGCTCAAGGTGGACACGCGGGAAGTTTGGCCCCCATTTCCGCGAGAG 729
Db      |||||
Qy 577 GCGGGCGCCTGTTCA----- 592
Db      |||||
Qy 730 GGAAACTCTACCTGATTGACTTGGCTGGGTGAGGACCAACCGCGGCACAGGCAACAAG 789
Db      |||||
Qy 593 -----TGATCGACCTGGCTGGCTCAGAGCGCGCTCGCAGACACAGNATCGT 639
Db      |||||
Qy 790 GGCCTTCGGTAAAGAGAGTGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTG 849
Db      |||||
Qy 640 GGCAGCGTATGAGGAGGGGGCCCATCAACCGCTCACTGTGGCACTGGGCAACTGC 699
Db      |||||
Qy 850 GTAGTGCCTGATCAG-----GGCCTCCCTCGTGTACTTATCGGACAGCAGCTC 903
Db      |||||
Qy 700 ATCAACGCTGTAGCGCAAGGGTAGCAACAAGTACATCAACTATCGCGACAGCAAGCTC 759
Db      |||||
Qy 904 ACTCGCCTATTGAGGAGCTCTTGGGTGGCTCAGGCCACAGTATCTTTATGGCAACATT 963
Db      |||||
Qy 760 ACCGGCTCTGAGGACTCTCTGGGAGGAAACAGCGCGCAGTATGATCGCTCATC 819
Db      |||||
Qy 964 GCCCTGAGAGAGCTTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCC 1023
Db      |||||
Qy 820 AGTCTTCGAGAGCTGCTTCGAGGAGTCCCGGAAACACCTGTACCTACGCCGCCGCCGCC 879
Db      |||||
Qy 1024 AAGGAGTGATCAATTG 1040
Db      |||||
Qy 880 AAGAACAATTAAAGACTAG 896
Db      |||||
```

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 19:47:28 ; Search time 4600.34 Seconds  
(without alignments)  
10701.074 Million cell updates/sec

Title: US-10-797-893-5  
Perfect score: 1041  
Sequence: 1 atgggtcgctgtcggtgaag.....ccaaggaggtgatcaattga 1041

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	100.0	1041	6	AR210054 Sequence
2	1038.4	99.8	1464	6	AR210055 Sequence
3	1038	99.7	1115	6	AR210052 Sequence
4	1035.4	99.5	1538	6	AR210053 Sequence
5	1035.4	99.5	1998	9	BT007259 Homo sapi
6	1035.4	99.5	1998	12	BT007888 Synthetic
7	1035.4	99.5	2097	6	AR304057 Sequence
8	1035.4	99.5	2097	9	AB017430 Homo sapi
9	1035.4	99.5	2117	9	BC028155 Homo sapi
10	1035.4	99.5	2134	9	BC004352 Homo sapi
11	772.8	74.2	2086	10	BC003427 Mus muscu
12	678	65.1	151041	2	AC101919 Mus muscu
13	678	65.1	164759	10	AC101752 Mus muscu
14	650.8	62.5	297639	2	AC101908 Mus muscu
15	522.8	50.2	196674	10	AC102127 Mus muscu
16	462.2	44.4	2756	5	BC063896 Xenopus t
17	454.6	43.7	2565	5	BC043733 Xenopus l
18	454.6	43.7	2583	5	XL249841 Xenopus l
19	454.6	43.7	2640	5	BC073177 Xenopus l

20	438.6	42.1	2593	5	AF267850 Xenopus l
21	438.6	42.1	2661	5	XL2429840 Xenopus l
22	438.6	42.1	2781	5	BC070549 Xenopus l
23	438.6	42.1	2819	5	AF267849 Xenopus l
24	355.4	34.1	174007	9	AC023831 Homo sapi
25	355.4	34.1	174363	9	AC009133 Homo sapi
26	352.2	33.8	3396	9	AB017333S2 Homo sapi
27	346.4	33.3	444	10	AF013119 Mus muscu
28	343.4	33.0	105485	9	HUAC002301 Homo sapi
29	278.4	26.7	68916	2	AC101927 Mus muscu
30	265	25.5	228783	2	AC094377 Rattus no
31	265	25.5	235806	2	AC133764 Rattus no
32	265	25.5	289504	2	AC123480 Rattus no
33	262.4	25.2	195987	10	AC122537 Mus muscu
34	262.4	25.2	226601	10	AC122863 Mus muscu
35	205.2	19.7	1982	3	AK114179 Ciona int
36	184.6	17.7	2095	6	AX833125 Sequence
37	184.6	17.7	2095	9	AK094619 Homo sapi
38	153.2	14.7	146415	5	AL929469 Zebrafish
39	153.2	14.7	176906	2	CR354587 Danio rer
40	151.8	14.6	3358	10	BC057614 Mus muscu
41	145.8	14.0	435	5	MSU34658 Morone saxa
42	145.2	13.9	44004	2	AC092563 Hylobates
43	143	13.7	183586	9	AC092562 Papio ham
44	136.6	13.1	1152	6	AR454169 Sequence
45	136.6	13.1	4108	6	AR454167 Sequence

ALIGNMENTS

RESULT 1  
LOCUS AR210054 1041 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 5 from patent US 6387644.  
ACCESSION AR210054  
VERSION AR210054.1 GI:21512185.  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1041)  
AUTHORS Beraud,C.  
TITLE Motor proteins and methods for their use  
JOURNAL Patent: US 6387644-A 5 14-MAY-2002;  
FEATURES  
source Location/Qualifiers  
1..1041  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match	100.0%	Score 1041;	DB 6;	Length 1041;
Best Local Similarity	100.0%;	Pred. No. 1.9e-267;		
Matches 1041;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGGTA	60	
Db	1	ATGGGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCGTCACCTCCAGCTCGGTA	60	
QY	61	AGGGTGGCTGTGCGACTGCGGCATTTGTGGATGGAAACAGCGGGAGCAAGTGTATCCCCC	120	
Db	61	AGGGTGGCTGTGCGACTGCGGCATTTGTGGATGGAAACAGCGGGAGCAAGTGTATCCCCC	120	
QY	121	TGTGTCGGGGGATGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGGACCAACCCAGGAG	180	
Db	121	TGTGTCGGGGGATGACAGCTGCTCTTAGAGATTGCTTAAGTGGAGGACCAACCCAGGAG	180	
QY	181	ACTCTCAATACACAGTTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTAT	240	
Db	181	ACTCTCAATACACAGTTTTCATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTAT	240	
QY	241	CGAGGTTTCTAGTCAGGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGCCTT	300	
Db	241	CGAGGTTTCTAGTCAGGCCCATCTTAAGGCACTTGTCTGGAAGGGCAGAAATGCCAGTGCCTT	300	

Db 241 GCAGGTTTCAGTCAGAGCCCATCTTAAGGCACTTCTGGAAGGCGAGAAATGCCAGTGTGCTT 300  
 QY 301 GCCTATGAGACCCACAGAGCTGGGAAGACGACACAATCTCTGGCGAGCCCGAGAGCAACT 360  
 Db 301 GCCTATGAGACCCACAGAGCTGGGAAGACGACACAATCTCTGGCGAGCCCGAGAGCAACT 360  
 QY 361 GGGGTGATCCCGGGGCTCTCATGAGACCTCTGAGACCTCTGAGCTCACAGGAGAGGGTCCGAG 420  
 Db 361 GGGGTGATCCCGGGGCTCTCATGAGACCTCTGAGACCTCTGAGCTCACAGGAGAGGGTCCGAG 420  
 QY 421 GCGCGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTTACAGAGAGGTA 480  
 Db 421 GCGCGCCATGGGCCCTTTCTGTACCATGTCTTACCTAGAGATCTTACAGAGAGGTA 480  
 QY 481 TTAGACCTCTGAGACCTCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGGTAAT 540  
 Db 481 TTAGACCTCTGAGACCTCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGGTAAT 540  
 QY 541 ATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCAC 600  
 Db 541 ATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCAC 600  
 QY 601 TTCTGCCAGCCAGTCCGAATCGGACTGTAGAGCCACCGGCTCAACAGAGCTCTCTCC 660  
 Db 601 TTCTGCCAGCCAGTCCGAATCGGACTGTAGAGCCACCGGCTCAACAGAGCTCTCTCC 660  
 QY 661 CCGAGTCATGCTGTCTCTGCTCAAGGTGAGAGCCAGCGGGAACGTTTGGCGGCAC 720  
 Db 661 CCGAGTCATGCTGTCTCTGCTCAAGGTGAGAGCCAGCGGGAACGTTTGGCGGCAC 720  
 QY 721 CAGCAGAGAGGGAACCTCTACCTGATTTGCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 780  
 Db 721 CAGCAGAGAGGGAACCTCTACCTGATTTGCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 780  
 QY 781 GGCACAAAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTG 840  
 Db 781 GGCACAAAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTG 840  
 QY 841 GGCACAAAGGTGATGCGCTGAATCAGGCTCTCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 900  
 Db 841 GGCACAAAGGTGATGCGCTGAATCAGGCTCTCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 900  
 QY 901 CTCACCTCGCTATTGAGAGCTCTCTGGTGGCTCAGCGCCACAGTATCTTATTGCCAAC 960  
 Db 901 CTCACCTCGCTATTGAGAGCTCTCTGGTGGCTCAGCGCCACAGTATCTTATTGCCAAC 960  
 QY 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020  
 Db 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020

RESULT 2  
 AR210055  
 LOCUS AR210055 1464 bp DNA linear PAT 20-JUN-2002  
 DEFINITION Sequence 7 from patent US 6387644.  
 ACCESSION AR210055  
 VERSION AR210055.1 GI:21512186  
 KEYWORDS

ORGANISM Unknown.  
 Unclassified.  
 Beraud,C.  
 TITLE Motor proteins and methods for their use  
 JOURNAL Patent: US 6387644-A 7 14-MAY-2002;  
 FEATURES  
 source  
 1..1464  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 99.8%; Score 1038.4; DB 6; Length 1464;  
 Best Local Similarity 99.9%; Pred. No. 9.5e-267;  
 Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGGGTCTCTGTGGCTAAGCAAGATTGGAGCTACTCTGTCGTCCACCTCCAGCTCCGCTA 60  
 Db 1 ATGGGTCTCTGTGGCTAAGCAAGATTGGAGCTACTCTGTCGTCCACCTCCAGCTCCGCTA 60  
 QY 61 AGGGTGGCTGTGGCTAGGAGCTGGGCTTGTGATGAGAAACAGCGGGAGCAAGTATCCCCC 120  
 Db 61 AGGGTGGCTGTGGCTAGGAGCTGGGCTTGTGATGAGAAACAGCGGGAGCAAGTATCCCCC 120  
 QY 121 TGTGTGGCGGCTAGGAGCAGCTGCTCTTAGAGATTGTCTAAGTGGAGAAACACAGGAG 180  
 Db 121 TGTGTGGCGGCTAGGAGCAGCTGCTCTTAGAGATTGTCTAAGTGGAGAAACACAGGAG 180  
 QY 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCTACAGAGACATCTAT 240  
 Db 181 ACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCTACAGAGACATCTAT 240  
 QY 241 GCAGTTTCAGTGCAGCCCATCTTAAGSCACTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTT 300  
 Db 241 GCAGTTTCAGTGCAGCCCATCTTAAGSCACTTGTCTGGAAGGGCAGAAATGCCAGTGTGCTT 300  
 QY 301 GCCTATGAGCCACAGGAGCTGGGAAGACGACACAATGTCTGGGAGCCCGAGAGCAACT 360  
 Db 301 GCCTATGAGCCACAGGAGCTGGGAAGACGACACAATGTCTGGGAGCCCGAGAGCAACT 360  
 QY 361 GGGGTGATCCCGGGGCTCTCATGAGCTCTGAGCTCAACAGGAGAGGGTGGCGAG 420  
 Db 361 GGGGTGATCCCGGGGCTCTCATGAGCTCTGAGCTCAACAGGAGAGGGTGGCGAG 420  
 QY 421 GCGCGGCTATGGGCTTCTGTCTCAACATGCTTCTACCTAGAGATCTACAGAGAGAGGTA 480  
 Db 421 GCGCGGCTATGGGCTTCTGTCTCAACATGCTTCTACCTAGAGATCTACAGAGAGAGGTA 480  
 QY 481 TTAGACCTCTGAGACCTCTGCTTGGGAGACCTGTAATCTCGAGAGAGCTGCGGGGGAAT 540  
 Db 481 TTAGACCTCTGAGACCTCTGCTTGGGAGACCTGTAATCTCGAGAGAGCTGCGGGGGAAT 540  
 QY 541 ATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTCTGATTTTGGAGCGGCAC 600  
 Db 541 ATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTCTGATTTTGGAGCGGCAC 600  
 QY 601 TTCTGCCAGCCAGTCCGAATCGGACTGTAGAGCCACCGGCTCAACAGAGCTCTCTCC 660  
 Db 601 TTCTGCCAGCCAGTCCGAATCGGACTGTAGAGCCACCGGCTCAACAGAGCTCTCTCC 660  
 QY 661 CCGAGTCATGCTGTCTCTGCTCAAGGTGAGAGCCAGCGGGAACGTTTGGCGGCAC 720  
 Db 661 CCGAGTCATGCTGTCTCTGCTCAAGGTGAGAGCCAGCGGGAACGTTTGGCGGCAC 720  
 QY 661 GGCACAAAGGTGATGCGCTGAATCAGGCTCTCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 780  
 Db 661 GGCACAAAGGTGATGCGCTGAATCAGGCTCTCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 780  
 QY 781 GGCACAAAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTG 840  
 Db 781 GGCACAAAGGGCTTGGCTAAAGAGAGTGGAGCCATCAACCTCTCTGTTTGTCTG 840  
 QY 841 GGCACAAAGGTGATGCGCTGAATCAGGCTCTCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 900  
 Db 841 GGCACAAAGGTGATGCGCTGAATCAGGCTCTCTGAGAGCCAGCGGGAACGTTTGGCGGCAC 900  
 QY 901 CTCACTCGCTATTGAGAGCTCTCTGGTGGCTCAGCGCCACAGTATCTTATTGCCAAC 960  
 Db 901 CTCACTCGCTATTGAGAGCTCTCTGGTGGCTCAGCGCCACAGTATCTTATTGCCAAC 960  
 QY 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020  
 Db 961 ATTGCCCCCTGAGAGAGCTTCTACCTAGACAGTCTCGGACTCAACTTTGCTGCCAGG 1020







QY 364 GTGATCCCGCGGCTCTCATGGACCTCTCTGAGCTCAAGGGAGGAGGGTGCAGAGGC 423  
 Db |||||  
 QY 436 GTGATCCCGCGGCTCTCATGGACCTCTCTGAGCTCAAGGGAGGAGGGTGCAGAGGC 495  
 Db |||||  
 QY 424 CGGCCATGGGCGCTCTCTGACCATCTTACCTAGAGATCTACACAGGAGAGGTATTA 483  
 Db |||||  
 QY 496 CGGCCATGGGCGCTCTCTGACCATCTTACCTAGAGATCTACACAGGAGAGGTATTA 555  
 Db |||||  
 QY 484 GACCTCTCGGACCTCTCTCGGAGACTGTGTAATCCGAGAGACTGCGGGGGAATATC 543  
 Db |||||  
 QY 556 GACCTCTCGGACCTCTCTCGGAGACTGTGTAATCCGAGAGACTGCGGGGGAATATC 615  
 Db |||||  
 QY 544 CTGATTCGGGCTCTCTCCGAGACCCATCAGTAGCTTTGCTGATTTTGGAGCGCATTC 603  
 Db |||||  
 QY 616 CTGATTCGGGCTCTCTCCGAGACCCATCAGTAGCTTTGCTGATTTTGGAGCGCATTC 675  
 Db |||||  
 QY 604 CTGCCACGAGTCCGAAATCGGACTGTAGGAGCCACCGGCTCAACACGAGCGCTCTCCCGC 663  
 Db |||||  
 QY 676 CTGCCACGAGTCCGAAATCGGACTGTAGGAGCCACCGGCTCAACACGAGCGCTCTCCCGC 735  
 Db |||||  
 QY 664 AGTCATCTGTCTCTCTGCTCAAGGTGGACCGAGCGGAAAGTCTTGGGCCCATTTTCGCCAG 723  
 Db |||||  
 QY 736 AGTCATCTGTCTCTCTGCTCAAGGTGGACCGAGCGGAAAGTCTTGGGCCCATTTTCGCCAG 795  
 Db |||||  
 QY 724 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTACAGAGCAACCGCGCGACAGGC 783  
 Db |||||  
 QY 796 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTACAGAGCAACCGCGCGACAGGC 855  
 Db |||||  
 QY 784 AACAGGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCTCTTGTCTTGGCTGGC 843  
 Db |||||  
 QY 856 AACAGGGGCTCTGGCTTAAAGAGAGTGGAGCCATCAACCTCTCTTGTCTTGGCTGGC 915  
 Db |||||  
 QY 844 AAAGTGTAGATGGCTGAATCAGGGCTCTCTGCTGACTTATCGGACAGCAAGCTC 903  
 Db |||||  
 QY 916 AAAGTGTAGATGGCTGAATCAGGGCTCTCTGCTGACTTATCGGACAGCAAGCTC 975  
 Db |||||  
 QY 904 ACTGCTTATGAGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATGCAACATT 963  
 Db |||||  
 QY 976 ACTGCTTATGAGAGCTCTCTGGGTGGCTCAGCCACAGTATCTTATGCAACATT 1035  
 Db |||||  
 QY 964 GCCCTGAGAGAGCTCTACCTAGACACAGTCTCCGACATCACTTGTGCGCAGTCC 1023  
 Db |||||  
 QY 1036 GCCCTGAGAGAGCTCTACCTAGACACAGTCTCCGACATCACTTGTGCGCAGTCC 1095  
 Db |||||  
 QY 1024 AAGGAGGTGATCAATTG 1040  
 Db |||||  
 QY 1096 AAGGAGGTGATCAATCG 1112  
 Db |||||

RESULT 6  
 BT007888  
 LOCUS 1998 bp mRNA linear SYN 13-MAY-2003  
 DEFINITION Synthetic construct Homo sapiens kinesin-like 4 mRNA, partial cds.  
 ACCESSION BT007888  
 VERSION BT007888.1 GI:30584614  
 KEYWORDS FLI CDNA.  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1 (bases 1 to 1998)  
 AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M., and Farmer,A.  
 TITLE Cloning of human full-length CDSs in BD Creator (TM) System Donor vector  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1998)  
 AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M., and Farmer,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

## COMMENT

This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion (TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

## FEATURES

## source

1..1998  
 /organism="synthetic construct"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:32630"  
 /clone="GH0024511.0"  
 /clone\_lib="BD Creator (TM) CDS Library derived from MOC collection"  
 /lab\_host="DH5alpha T1 resistant"  
 /note="Vector: pDNR-Dual"

## CDS

1..>1998  
 /note="Mutations: 1997:Stop->Leu"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Homo sapiens kinesin-like 4"  
 /protein\_id="AAP36560.1"  
 /db\_xref="GI:30584615"  
 /translation="MAAGSTQORRREMAAASAAISGACRCLSKIGATRRPPPARV  
 RVAVRPFDVTAGADPPCVRGMSDSLEIANWRNHQHTLTKYQFADPYGERSTQDD  
 IVAGSVQPIILHLLGONASVLAAYGTAGTKTILMGSPQPGVIPALMDLLQLTRE  
 ADAGRPWALSVTMSYLEIQEVKLDLDPASGDVIRDCRGNILIPGISQKPISSF  
 EAPERHFLPASNRVTGALRNORSRSHVLLVKVDQRELRAPRFQRECKLYLIDLA  
 GSDNRRTNGKGLRLKESGAINTSFLVGVKVDALNQGRLPVPRDSKLRLLQDSLG  
 GSAHSLIANIAPERPYLDTVSALNPAARSKEVINRPTNESLQPHALGPKVLSQKE  
 LIGPPAKARAGPEEIEGSPGMAAPASAKSLPLQKLSMDPAMWLEKLSLDRL  
 ASGSGQAPLLSTPKRVMVLMKTVBEKDLLEIRLTKQKEAKMLAQAEKENHC  
 PTMLRSLRHTVTGPKLKKAVVMPLQLIQEQAASPAEIHILNKIKRKKKLSLDAL  
 EPEKAEDQWELQISPELLAHGROKILDLINEGSARDLRSIORGPKKQLIVGWREL  
 HGPFQSVEDLERVEGITGQMESFLKNIILGLAAGQRCGASL"

## ORIGIN

Query Match 99.5%; Score 1035.4; DB 12; Length 1998;  
 Best Local Similarity 99.9%; Pred. No. 6.1e-266;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 63  
 Db |||||  
 QY 76 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGGTAAGG 135  
 Db |||||  
 QY 64 GTGCTGTGCGCTGCGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCCCTGT 123  
 Db |||||  
 QY 136 GTGCTGTGCGCTGCGGCCATTTGTGGATGGAAACAGCGGAGCAAGTGTATCCCCCTGT 195  
 Db |||||  
 QY 124 GTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 183  
 Db |||||  
 QY 196 GTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 255  
 Db |||||  
 QY 184 CTCAAATACAGTTTTCATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243  
 Db |||||  
 QY 256 CTCAAATACAGTTTTCATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 315  
 Db |||||  
 QY 244 GGTTCAGTGCAGGCCATCTTAAGGCATTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGGC 303  
 Db |||||  
 QY 316 GGTTCAGTGCAGGCCATCTTAAGGCATTTGCTGGAAGGGCAGAAATGCCAGTGTGCTTGGC 375  
 Db |||||  
 QY 304 TATGGAACCAAGAGCTGGGAAGACGACACAATGCTGGGACGCCAGAGCAACCTGGG 363  
 Db |||||  
 QY 376 TATGGAACCAAGAGCTGGGAAGACGACACAATGCTGGGACGCCAGAGCAACCTGGG 435  
 Db |||||  
 QY 364 GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGC 423  
 Db |||||  
 QY 436 GTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGC 495  
 Db |||||

QY	424	CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCGAGAGAGGTATTA	483
DB	496	CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCGAGAGAGGTATTA	555
QY	484	GACCTCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCGGGGGGAATATC	543
DB	556	GACCTCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCGGGGGGAATATC	615
QY	544	CTGATTCCGGGCTCTCCAGAAAGCCCATCAGTAGCTTTGTGATTTTGGAGCGGCATTC	603
DB	616	CTGATTCCGGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTGTGATTTTGGAGCGGCATTC	675
QY	604	CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGGCTCTCCCGC	663
DB	676	CTGCCAGCCAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGGCTCTCCCGC	735
QY	664	AGTCATGCTGCTCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTCCGCCAG	723
DB	736	AGTCATGCTGCTCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTCCGCCAG	795
QY	724	CGAGAGGGAATACTCTACCTGATTGACTTGGCTGGGTGAGGACAACCGGCGCACAGGC	783
DB	796	CGAGAGGGAATACTCTACCTGATTGACTTGGCTGGGTGAGGACAACCGGCGCACAGGC	855
QY	784	ACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCTGGGC	843
DB	856	ACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCTGGGC	915
QY	844	AAAGTGTAGATCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGAGACGAAGCTC	903
DB	916	AAAGTGTAGATCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGAGACGAAGCTC	975
QY	904	ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGGCCACAGATCTTATTCGCAACATT	963
DB	976	ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGGCCACAGATCTTATTCGCAACATT	1035
QY	964	GCCCTGAGAGAGCTTCTACCTAGACAGTCTCGCACTCAACTTTGCTGCCAGGTCC	1023
DB	1036	GCCCTGAGAGAGCTTCTACCTAGACAGTCTCGCACTCAACTTTGCTGCCAGGTCC	1095
QY	1024	AAGGAGGTGATCAATTG 1040	
DB	1096	AAGGAGGTGATCAATCG 1112	
RESULT 7			
AR304057			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1036; Conservative			
QY			
DB			

QY	64	GTGGCTGTGCGACTGCGGCCATTTGTGATCGAACAGCGGAGCAAGTGATCCCCCTGT	123
DB	160	GTGGCTGTGCGACTGCGGCCATTTGTGATCGAACAGCGGAGCAAGTGATCCCCCTGT	219
QY	124	GTGCGGGCATGAGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGAGACT	183
DB	220	GTGCGGGCATGAGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGAGACT	279
QY	184	CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGAGGACATCTATGCA	243
DB	280	CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGAGGACATCTATGCA	339
QY	244	GGTTCACTGACAGCCATCTTAAGCACTTGTCTGAGAGGAGGAGTATGCGAGTGTCTGCC	303
DB	340	GGTTCACTGACAGCCATCTTAAGCACTTGTCTGAGAGGAGGAGTATGCGAGTGTCTGCC	399
QY	304	TATGACCCACAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGAGCAACTGGG	363
DB	400	TATGACCCACAGAGAGCTGGGAAGACGACACAATGCTGGGAGCCACAGAGCAACTGGG	459
QY	364	GTGATCCCGCGGCTCTCATGAGCCTCTCTGAGCTCAACAGGAGAGGGTGCAGAGGC	423
DB	460	GTGATCCCGCGGCTCTCATGAGCCTCTCTGAGCTCAACAGGAGAGGGTGCAGAGGC	519
QY	424	CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCGAGAGAGGTATTA	483
DB	520	CGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCGAGAGAGGTATTA	579
QY	484	GACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCGGGGGAATATC	543
DB	580	GACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCGGGGGAATATC	639
QY	544	CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGGCGGCACATTC	603
DB	640	CTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGGCGGCACATTC	699
QY	604	CTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCCCGC	663
DB	700	CTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCCCGC	759
QY	664	AGTCATGCTGCTCTCTGCTCAAGGTGACAGCGGGAAACGTTTGGCCCCATTTGCCAG	723
DB	760	AGTCATGCTGCTCTCTGCTCAAGGTGACAGCGGGAAACGTTTGGCCCCATTTGCCAG	819
QY	724	CGAGAGGAAATCTACTGATTGACTTGGCTGGGTGAGGAGCAACCGGCGCACAGGC	783
DB	820	CGAGAGGAAATCTACTGATTGACTTGGCTGGGTGAGGAGCAACCGGCGCACAGGC	879
QY	784	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTCTGGGC	843
DB	880	AACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCTGTTTGTCTCTGGGC	939
QY	844	AAAGTGTGATGCGCTGAATCAGGGCTCTGCTGATACCTTATCGGAGAGAGCTC	903
DB	940	AAAGTGTGATGCGCTGAATCAGGGCTCTGCTGATACCTTATCGGAGAGAGCTC	999
QY	904	ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGGCCACAGATCTTATTCGCAACATT	963
DB	1000	ACTGCGCTATTGAGAGACTCTCTGGGTGGCTCAGGCCACAGATCTTATTCGCAACATT	1059
QY	964	GCCCTGAGAGAGCTTCTACCTAGACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC	1023
DB	1060	GCCCTGAGAGAGCTTCTACCTAGACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC	1119
QY	1024	AAGGAGGTGATCAATTG 1040	
DB	1120	AAGGAGGTGATCAATCG 1136	
RESULT 8			
AB017430			
LOCUS			
Query Match			
Best Local Similarity			
Matches 1036; Conservative			
QY			
DB			

AB017430  
LOCUS

2097 bp mRNA linear PRI 06-MAR-1999

DEFINITION Homo sapiens mRNA for kinesin-like DNA binding protein, complete cds.

ACCESSION AB017430 D38751

VERSION AB017430.2 GI:4519442

KEYWORDS Kid; kinesin-like DNA binding protein; kinesin family.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (sites)

AUTHORS Tokai N., Fujimoto-Nishiyama A., Toyoshima Y., Yonemura S., Tsukita S., Inoue J. and Yamamoto T.

TITLE Kid, a novel kinesin-like DNA binding protein, is localized to chromosomes and the mitotic spindle

JOURNAL ENBO J. 15 (3), 457-467 (1996)

MEDLINE 96174806

PUBMED 8599929

REFERENCE 2 (bases 1 to 2097)

AUTHORS Tokai Nishizumi, N. and Edamasa, M.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-1998) Noriko Tokai Nishizumi, The University of Tokyo, The Institute of Medical Science, Department of Oncology; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan (E-mail: tokai@hgc.ims.u-tokyo.ac.jp, Tel: 03-5449-5302, Fax: 03-5449-5413)

COMMENT D38751: Submitted (02-Nov-1994)

Sequence updated (25-Feb-1999).

FEATURES

Location/Qualifiers

1..2097

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="pBSkid"

/cell\_line="MDA-MB453"

/cell\_type="breast cancer"

1..2097

/genes="Kid"

25..2022

/genes="Kid"

/codon\_start=1

/product="kinesin-like DNA binding protein"

/protein\_id="BA033019.2"

/db\_xref="GI:4519443"

/translation="MAAGSTQQRREMAAASAAISGAGRCRLSKIGATRRPPARV  
RYVRLRPVFDGAGSDPCVCGMDSCSLEIANWNRHOETLKYOFDAFYGRSTOOD  
IVAGSVQPIRLHLEQNASVLAYGPTGAGKTHMLGSPEQGVIPRALMDLLQTLRE  
EGAGRPWALSTMVSLYIEYQKVLDPASGLVIREDCRGNILIPGLOSOKPISFF  
ADPERFLPASNRRTVGTALNRQSRSHAVLLVVDQRLAPRQREGKLYLDLA  
GSDNRNRNKGRLKESGAINTSFLVKGVDALNQGLRPVPRYDSKLIRLLOPSLG  
GSAHSILIANIAPRRFYDITVSNLFAASKEVINRPTNESLOPHALGPVKSQKE  
LLGPPKARGPPEEIEIGSPREMAAPASAKLSPLQKLSMDPAMLERLISLRLLL  
ASQSGQAPLLSTPKRERMVLMKTVEEKDLIERLTKQKELEAKMLAQKAEKENHC  
PTMLRPLSRITVGAKPLKKAVMVPLQIQSQNASPNABIHILKNRGRKRLSDAL  
EPEEKADCEWELQISPELLAHGRQKLLDNLNFGSARDLSLQRIQPKKAQLIVGWREL  
HGPFQVEDLERVEGITGKMFSLKANTILGLAAGRCGAS"

ORIGIN

Query Match 99.5%; Score 1035.4; DB 9; Length 2097;

Best Local Similarity 99.9%; Pred. No. 6.1e-266;

Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCGTGTGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63

DB 100 GGTGCGTGTGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 159

QY 64 GTGGCTGTGGCTCGGCGCATTTGTGGATGGAACAGCGGGAGCAAGTATGATCCCCCTGT 123

DB 160 GTGGCTGTGGCTCGGCGCATTTGTGGATGGAACAGCGGGAGCAAGTATGATCCCCCTGT 219

QY 124 GTGGGGGGGATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGACCAACAGGAGACT 183

DB 220 GTGGGGGGGATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGACCAACAGGAGACT 279

QY 184 CTCAATATACAGTTTGATGCTTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243

DB 280 CTCAATATACAGTTTGATGCTTTCTATGGGGAGAGGAGTACTCAGCAGGACATCTATGCA 339

QY 244 GGTTCAGTCAGGCCCATCTTAAGGCACTTCTCTGGAAGGGCAGAAATGCAAGTGTCTTGGC 303

DB 340 GGTTCAGTCAGGCCCATCTTAAGGCACTTCTCTGGAAGGGCAGAAATGCAAGTGTCTTGGC 399

QY 304 TATGGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGCGAGCCCGAGAGCAACCTGGG 363

DB 400 TATGGACCCACAGGAGCTGGGAAGAGCGACACAAATGCTGGGCGAGCCCGAGAGCAACCTGGG 459

QY 364 GTGATCCCGGGGCTCTCATGAGCACTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 423

DB 460 GTGATCCCGGGGCTCTCATGAGCACTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 519

QY 424 CGGCCATGGGCGCTTCTGTACCATGTCTTACTAGAGATCTTACAGGAGAGAGGATTA 483

DB 520 CGGCCATGGGCGCTTCTGTACCATGTCTTACTAGAGATCTTACAGGAGAGAGGATTA 579

QY 484 GACCTCTGAGCCCTCTCTCGGGAGACCTGGTAATCCGAGAAGCACTCCCGGGGGAATATC 543

DB 580 GACCTCTGAGCCCTCTCTCGGGAGACCTGGTAATCCGAGAAGCACTCCCGGGGGAATATC 639

QY 544 CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTTGTCTGATTTTGAAGCGGCACTTC 603

DB 640 CTGATTCGGGGTCTCTCCAGAGCCCATCAGTAGCTTTTGTCTGATTTTGAAGCGGCACTTC 699

QY 604 CTGCGCAGCAGTCGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGGCGCTCTCTCCCGC 663

DB 700 CTGCGCAGCAGTCGAAATCGGACTGTAGGAGCCACCGGGCTCAACAGGCGCTCTCTCCCGC 759

QY 664 AGTCATGCTGCTCTCTGCTCAAGGTGGACCGGGGAGCGTTTGGCCCATTTCCGCCAG 723

DB 760 AGTCATGCTGCTCTCTGCTCAAGGTGGACCGGGGAGCGTTTGGCCCATTTCCGCCAG 819

QY 724 CGAGAGGGGAAACTCTTACCTGATTTGACTTGGCTGGGTGAGAGGACAAACCGGCGCACAGGC 783

DB 820 CGAGAGGGGAAACTCTTACCTGATTTGACTTGGCTGGGTGAGAGGACAAACCGGCGCACAGGC 879

QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCTCTGTTTGTCTGGGC 843

DB 880 AACAGGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCTCTGTTTGTCTGGGC 939

QY 844 AAGTGGTAGAGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 903

DB 940 AAGTGGTAGAGCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 999

QY 904 ACTCGCTATTGCGAGGACTCTCTGGGTGGCTCAGGCCACAGCATCTTATTTGCCAACATT 963

DB 1000 ACTCGCTATTGCGAGGACTCTCTGGGTGGCTCAGGCCACAGCATCTTATTTGCCAACATT 1059

QY 964 GCGCTGTAGAGAGCGTTTCTTACCTAGACAGTCTCCGCACTCAACTTTGCTCCAGGTC 1023

DB 1060 GCGCTGTAGAGAGCGTTTCTTACCTAGACAGTCTCCGCACTCAACTTTGCTCCAGGTC 1119

QY 1024 AAGGAGGTGATCAATTG 1040

DB 1120 AAGGAGGTGATCAATCG 1136

RESULT 9

LOCUS BC028155

DEFINITION Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:40049 IMAGE:5241557), complete cds.

ACCESSION BC028155

VERSION BC028155.1 GI:20380446

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 2117)  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,  
Sheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,  
Carinci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,  
Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,  
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettner M., Madan A., Rodriguez S.,  
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,  
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,  
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E.,  
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE

JOURNAL  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

2 (bases 1 to 2117)  
Strausberg R.  
Direct Submission  
Submitted (08-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nigri.nih.gov](mailto:nisc_mgc@nigri.nih.gov)  
Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,  
Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,  
Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P.,  
Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,  
Maduro Q.L., Masiello C., Maskeri B., Mastrian S.D., McCloskey J.C.,  
McDowell J., Pearson R., Stantripop S., Thomas P.J., Touchman J.W.,  
Tsurgoun C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,  
Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Series: IRAC Plate: 62 Row: 1 Column: 3  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, Similarity but not identity to protein.

FEATURES

source

1..2117  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:40049 IMAGE:5241557"  
/tissue\_type="Brain, fetal, whole pooled"  
/clone\_lib="NIH MGC\_121"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

gene

1..2117  
/gene="KIF22"  
/note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"  
/db\_xref="LocusID:3835"  
/db\_xref="MIM:603213"

CDS

6..2003  
/gene="KIF22"

/codon\_start=1  
/product="kinesin family member 22"  
/protein\_id="AAH28155.1"  
/db\_xref="GI:20380447"  
/db\_xref="LocusID:3835"  
/db\_xref="MIM:603213"  
/translation="MAAGGSGTQORREMAAASAAIAGACRLSKIGATRRPPARV  
RVARLPFFVGTGASDPCVGMDSLEIANRNHETLKQFDYAFGERSTQDD  
IYAGSDPILRHLLGGQNASVLYGTGAGKTHMLGSPQGVIPRALMDLQITRE  
EGAEKFWALSVTMSYLEIYQKVDLDLDPSGDLVREDCRGNILIPGLSQPTSSF  
ADFERHFLPASNRNRTVGATRNQSSRSRSHVLLVKVDORERLAPROREGKLYLIDLA  
GSENRRTNGKRLKSGSANTSLFVLGVKVDALNOGLPRVPYVDSKLTPLQDLSLG  
GSAHSILLIANTAPERREYFVTSALNFAARSKVEINRPTNESLQPHALGPVLSQKE  
LIGPEAKRGRPEEERIGSPMAAPASAKLSPQKSPQSSMDPAMLERLSLDRLL  
ASQSGAPLSTTPKRRKVMIMTKTEERLTKQKLEKRLAKMLAKAEKHEHC  
PTMLRPLSHLTVTGAKPLKAVVMPLOIQPAASPNRATIHILNKKRKRKGLSDAL  
EPEKAEKDCWELQISPELLAHGROKILDLNLESGARDLSRLQIPKPKAQLIVGWREL  
HGPFQVEDLERVSGITGKQWESFLKANILGLAAGRCGAS"

ORIGIN

Query Match 99.5%; Score 1035.4; DB 9; Length 2117;  
Best Local Similarity 99.9%; Pred. No. 6.1e-266;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 GGTGCTGTCGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63  
DB 81 GGTGCTGTCGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG 140  
QY 64 GTGGCTGTGCACTGCGGCCAATTTGTGATGGAACACAGCGGGAGCAAGTGATCCCCCTGT 123  
DB 141 GTGGCTGTGCACTGCGGCCAATTTGTGATGGAACACAGCGGGAGCAAGTGATCCCCCTGT 200  
QY 124 GTGGCGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGAGAGAACACACAGAGACT 183  
DB 201 GTGGCGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGAGAGAACACACAGAGACT 260  
QY 184 CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCTAGAGAGATCTATGCA 243  
DB 261 CTCAAAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCTAGAGAGATCTATGCA 320  
QY 244 GGTTCAGTGTGACGCCATCTTAAGGCACTTGTGGAAGGCGAGAATGCGAGTGTGCTGCC 303  
DB 321 GGTTCAGTGTGACGCCATCTTAAGGCACTTGTGGAAGGCGAGAATGCGAGTGTGCTGCC 380  
QY 304 TATGACCCACAGAGCTGGGAACACACACATGCTGGGACGCCACACAGACCTGGG 363  
DB 381 TATGACCCACAGAGCTGGGAACACACACATGCTGGGACGCCACACAGACCTGGG 440  
QY 364 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCAACAGGAGAGGGTGGCGAGGCG 423  
DB 441 GTGATCCCGCGGCTCTCATGACCTCTGAGCTCAACAGGAGAGGGTGGCGAGGCG 500  
QY 424 CGGCGATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACAGGAGAGAGTATTA 483  
DB 501 CGGCGATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACAGGAGAGAGTATTA 560  
QY 484 GACCTCTGACCCCTGCTTGGGAGACTGTGTAATCCGAGAGACTGCGGGGGAATATC 543  
DB 561 GACCTCTGACCCCTGCTTGGGAGACTGTGTAATCCGAGAGACTGCGGGGGAATATC 620  
QY 544 CTGATTCGGGCTCTCTCCAGAACCCATCAGTAGCTTTGTGATTTTGAGCGGCACTTC 603  
DB 621 CTGATTCGGGCTCTCTCCAGAACCCATCAGTAGCTTTGTGATTTTGAGCGGCACTTC 680  
QY 604 CTGCGACCGAGTCGAAATCGGACTGTAGAGGCCACCCGGCTCAACACAGCGCTCTCCGCG 663  
DB 681 CTGCGACCGAGTCGAAATCGGACTGTAGAGGCCACCCGGCTCAACACAGCGCTCTCCGCG 740  
QY 664 AGTCATGCTGTGCTCTGCTCAAGCTGACCGCGGGAACGTTTGGCCCAATTTGGCCAG 723  
DB 741 AGTCATGCTGTGCTCTGCTCAAGCTGACCGCGGGAACGTTTGGCCCAATTTGGCCAG 800  
QY 724 CGAGAGGGAATACTCTACCTGATTGACTTGGCTGGGTGAGAGGAACAACCGCGCACAGGC 783

```

|||||
801 CGAGAGGGAACCTCTACCTGATTTGCTTGGCTGAGGACCAACCGCGCAGGC 860
|||||
784 AACACGGGCTTCGGCTAAAGAGAGCTGGAGGCATCAACACCTTCCCTGTTGTCCTGGGC 843
|||||
861 AACACGGGCTTCGGCTAAAGAGAGCTGGAGGCATCAACACCTTCCCTGTTGTCCTGGGC 920
|||||
844 AAAGTGATAGTGGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903
|||||
921 AAAGTGATAGTGGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 980
|||||
904 ACTCGCTATTGACAGACTCTCTGGTGGCTAGCCACAGATATCTTATGCAACATT 963
|||||
981 ACTCGCTATTGACAGACTCTCTGGTGGCTAGCCACAGATATCTTATGCAACATT 1040
|||||
964 GCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTCTGCGAGGTCC 1023
|||||
1041 GCCCTCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTCTGCGAGGTCC 1100
|||||
1024 AAGGAGGTGATCAATTG 1040
|||||
1101 AAGGAGGTGATCAATCG 1117
|||||

BC004352      2134 bp  mRNA  linear  PRI 29-JUN-2004
Homo sapiens kinesin family member 22, mRNA (cDNA clone MGC:1573
IMAGE:3535435), complete cds.
BC004352
VERSION      BC004352.1  GI:13279307
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2134)
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalieu,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2134)
Straussberg,R.
Direct Submission
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,

FEATURES
source
1..2134
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:1573 IMAGE:3535435"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2134
/gene="KIF22"
/note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
20..2017
/gene="KIF22"
/codon_start=1
/product="kinesin family member 22"
/protein_id="AAH04352.1"
/db_xref="GI:13279308"
/db_xref="LocusID:3835"
/db_xref="MIM:603213"
/translacion="MAAGSGTQORREMAASAAISGAGRCLSKIGATRRPPPARV
RVAVLRFPVGTAGAPPCVGRMGDSLEIANRWNETKLYQFADYFGRSTQOD
IYAGVSPILRLHLEGQNASVLAYGTGAGTHTWLGSPQGVIPRALMDLLQLTRE
EGAGRPWALSVMYSLEYQKVDLLDPASGLVIREDCRNLILPGLSQKPTSSF
ADFERHFLPASNRNVTGATRLNQRSRSHAVLLVKVDORERLAPRQREGKLYIDL
GSEDNRRTNGKRLKESGAINTSFLVLGVVDALNOGLPRVYRDSKLTLLQDSLG
CSAHSILIANIAPERFYLDTVSALNFAARKEVINRPTNESLQPHALGPVKLSQKE
LIGPEAKRAGPEEEIGSPFMAAPASASOKISFLKLSMDPAMLERLLSLDKILL
ASQSGQAPLJSTPKRERMVLMKTVVEKDLEIRLTKQKLEAKMLAQKBEKENHC
PTMLRPLSHRTVTGAKPLKCAVMPLOIQEQASPNABEHLKNGRKRKLESIDL
EPEKAEDCWELOISPPELLAHGRQKILDLNNEGSRDLRLSLQRIQPKKQLIVGWREL
HGPFQSVEDLERVEGITGKQMESFLKANILGLAAGRCGAS"

ORIGIN
Query Match 99.5%; Score 1035.4; DB 9; Length 2134;
Best Local Similarity 99.9%; Pred. No. 6.1e-266;
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GGTGCTCTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCCGCTAAGG 63
DB 95 GGTGCTCTCGGCTAAGCAAGATTGGAGCTACTCGTCCACCTCCAGCTCCGCTAAGG 154
QY 64 GTGGCTGTGCGACTCGGCGCATTTTGTGATGGAACACGCGGAGCAAGTATCCCCCTGT 123
DB 155 GTGGCTGTGCGACTCGGCGCATTTTGTGATGGAACACGCGGAGCAAGTATCCCCCTGT 214
QY 124 GTGCGGGCATGACGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAACACCAAGAGACT 183
DB 215 GTGCGGGCATGACGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAACACCAAGAGACT 274
QY 184 CTCGAATACCAAGTTTGTGATGCTTCTTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243
DB 275 CTCGAATACCAAGTTTGTGATGCTTCTTATGGGAGAGAGTACTCAGCAGACATCTATGCA 334
```











## RESULT 13

AC101752/c  
 LOCUS AC101752 164759 bp DNA linear ROD 23-APR-2004  
 DEFINITION Mus musculus chromosome 1, clone RP24-344C18, complete sequence.  
 ACCESSION AC101752  
 VERSION AC101752.14 GI:46518680  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

## REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## AUTHORS

1 (bases 1 to 164759)  
 Birren, B., Nussbaum, C. and Lander, E.

## TITLE

Mus musculus chromosome 1, clone RP24-344C18

## REFERENCE

2 (bases 1 to 164759)  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,

## AUTHORS

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,  
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
 Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A.,  
 Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hago, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
 Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
 Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE

3 (bases 1 to 164759)  
 Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,  
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,  
 Collamore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K.,  
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,  
 Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,  
 Graham, L., Grand-Pierre, N., Hafez, N., Hago, P., Hago, B.,  
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,  
 Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
 Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,  
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
 Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,  
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE

4 (bases 1 to 164759)  
 Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,  
 Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V.,  
 Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J.,  
 Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B.,  
 DeArelano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,  
 Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D.,  
 Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,

Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,  
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,  
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,  
 Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,  
 McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,  
 Murphy, T., Naylor, J., Nguyen, C., O'Neill, D., Oliver, J., Peterson, K.,  
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
 Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,  
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R.,  
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,  
 Zimmer, A. and Zody, M.

## TITLE

Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Apr 23, 2004 this sequence version replaced gi:45504330.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

## -----

Center: Whitehead Institute/MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu

## -----

Contact: sequence submissions@broad.mit.edu  
 ----- Project Information  
 Center project name: L17344  
 Center clone name: 344\_C\_18

## -----

FEATURES  
 source  
 1..164759  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="1"  
 /map="1"  
 /clone="RP24-344C18"  
 /clone\_lib="RPC1-24 Male Mouse BAC"

## repeat\_region

1..1655  
 /rpt\_family="L1\_MM"

## misc\_feature

1..4  
 /note="clone boundary  
 clone end:SP6  
 site:Mbol"

## repeat\_region

1767..1787  
 /rpt\_family="AT\_rich"

## repeat\_region

2115..2157  
 /rpt\_family="AT\_rich"

## repeat\_region

2236..2258  
 /rpt\_family="AT\_rich"

## repeat\_region

3252..3420  
 /rpt\_family="Lx7"

## repeat\_region

3464..3588  
 /rpt\_family="Lx7"

## repeat\_region

3605..3666  
 /rpt\_family="CT-rich"

## repeat\_region

3673..3731  
 /rpt\_family="(TTC)n"

## repeat\_region

3732..3911  
 /rpt\_family="(TTC)n"

## repeat\_region

4000..4180  
 /rpt\_family="CT-rich"

## repeat\_region

4202..4363  
 /rpt\_family="Lx7"

## repeat\_region

4529..4765  
 /rpt\_family="MTE"

## repeat\_region

7234..7442  
 /rpt\_family="B3"

## repeat\_region

8445..8509  
 /rpt\_family="T-rich"

## repeat\_region

complement(8512..8591)  
 /rpt\_family="MLT1"

Query Match	65.11%	Score 678;	DB 10;	Length 164759;
Best Local Similarity	80.0%;	Pred. No. 2.8e-170;		
Matches	824;	Conservative 0;	Mismatches 200;	Indels 6; Gaps 2;
Qy	12	TCGCGTAAACAAGATTGGAGCTACTGTCGTCCACTCCAGCTCGCGTAAAGGTGGCTGT	71	
Db	146111	TTGTATAAGCAAGGTAGGATTGGCTGTCGTCTTCTAGCTCGAGTAAGGTGGCTAT	146052	
Qy	72	CGGACTGGCGGCATTGTGGATTGGAACAGCGGGAGCAAGTGAATCCCCCTGTGTGGGGG	131	
Db	146051	ATAATTGTGTCCTGTTGTGGATTGAAGAGACAGAACGGAAGGAGCACCCCTGTGTACGAGC	145992	
Qy	132	CATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACCAACGAGACCTCTCAATA	191	
Db	145991	CATAGACAGCTGCTCTCTTTGAATGGCTTAACCTGGAGGAAATACCAAGGACCTATCAAAAT	145932	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 297639)  
Birren,B., Nusbaum,C. and Lander,E.  
Mus musculus chromosome 1, clone RP24-275J1  
Unpublished  
2 (bases 1 to 297639)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,  
Choepe,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,  
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 297639)  
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,  
Choepe,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,  
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,  
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 27, 2004 this sequence version replaced gi:31880232.  
All repeats were identified using RepeatMasker:  
Smith,A.P.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@broad.mit.edu  
----- Project Information  
Center project name: L17732  
Center clone name: 275\_J\_1  
-----

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

ORIGIN

Query Match 62.5%; Score 650.8; DB 2; Length 297639;  
Best Local Similarity 78.3%; Pred. No. 5.3e-163;  
Matches 807; Conservative 0; Mismatches 217; Indels 6; Gaps 2;

QY 12 TCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGGTTGGCTGT 71  
Db 1179 TTGTATAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGGTTGGCTGT 1120  
QY 72 GCGACTCGCGCCATTGTGGATGACACGGGAGCAAGTGTATCCCTCTGTGCGGG 131  
Db 1119 ATATTTGTGCTTTTGTGGATGACACGGGAGCAAGTGTATCCCTCTGTGCGGG 1060  
QY 132 CATGGACAGCTGCTCTTAGAGATTCTTAACCTGGAGAACCCACAGAGACTCTCAATA 191  
Db 1059 CATAGACAGCTGCTCTTTGAAGTGGCTAACTGGAGNAATACACAGAGACTATCAATA 1000  
QY 192 CAGTTTGTATGCTTCTTATGGGAGAGAGTACTACAGAGGACATCTATGAGTTTCACT 251  
Db 999 TCAGTTTGTATGCTTCTTATGGGAGAGAGTACTACAGAGGACATCTATGAGTTTCACT 940  
QY 252 GCAGCCCTCTTAAGGCACTTCTGGAAGGCAAGTGTATCCCTCTGTGCGGG 311  
Db 939 ACAGCCCTCTTAAGGCACTTCTGGAAGGCAAGTGTATCCCTCTGTGCGGG 880  
QY 312 CACAGGAGCTGGAGAGACGACACATGCTGGGAGCCACAGAGCAACCTGGGTTGATCCC 371  
Db 879 GACTGGGAGGAG 822  
QY 372 GCGGGCTCTCATGGACCTCTCTGAGCTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 431  
Db 821 TCAGACTCTCCAGAGACCTCTCTGACCTCAACAGGAGAGAGAGAGAGAGAGAGAGAGAG 762  
QY 432 GCGCCCTCTCTGAGCAGCTCTTACCTAGAGATCTACAGAGAGAGAGAGAGAGAGAGAGAG 491  
Db 761 GGACATTTCTGTGCTTTTGTGCTTTTGTGCTTTTGTGCTTTTGTGCTTTTGTGCTTTT 702  
QY 492 GGACCTCTCTGCGGAGAGCTGCTATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551  
Db 701 GGATCTCTCTATAGTGTACCTAAATGTGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 642  
QY 552 GGGTCTCTCCAG 611  
Db 641 AGGCTCTCACAG 582

FEATURES  
source

1..297639  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clone="RP24-275J1"  
/clone\_lib="RP24-275J1"  
/clone="RP24-275J1"

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.

```
QY 612 CAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACGAGCGCTCTCCCGAGTCATGC 671
|||||
Db 581 CAGTAGAATCAGTGTAGGAGCTACTCTGCTTAAACGAGCGCTCTTCCCGTAGTCATGC 522
|||||
QY 672 TGTGCTCTCTGCTCAAGGTGGACCGGGAAGCTTTTGGCCCCCATTTCCCGCAGCGAGAGG 731
|||||
Db 521 AGCGCTCTTGGTCAGAGTAGATCAGCATGAGCTTTGACTCCATTTCCCGAGTGGGAAG 462
|||||
QY 732 AAACTCTACCTGATGATCTGGCTGGGTGAGAGCAACCC-----GGCGCAGGCAACA 787
|||||
Db 461 AAACTCTACCTTATGATTTGGCTGGTTCAGAGCAACCGCTCAGTCGCGCAGGCAACC 402
|||||
QY 788 AGGCGCTTCGGCTAAAGAGAGTGAGGCATCAACACCTCCCTGTTCTCTGGGCAAG 847
|||||
Db 401 AGGCGATCTCTCAAGAGAGTGAGGCATCAACACCTCCCTCTTTGTACTGGGCAAG 342
|||||
QY 848 TGGTAGATCGCGTGAATCAGGCGCTCCCTCGTGTAACCTTATCGGACAGCAAGTCACTC 907
|||||
Db 341 TAGTGATGATTAACCCAGGCTCCCTCATATACCATACAGGACAGGAGCTCACTC 282
|||||
QY 908 GCCTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCCAATGTCGCC 967
|||||
Db 281 GTCCTTGCAAGACTCTCTGAGAGGCTCAGGCTCATAGCATCTCATTTGTCAGCATGCTC 222
|||||
QY 968 CTGAGAGCGCTTCTACCTAGACAGCTCGGCACTCAACTCTGCTCCGAGTCCAAAG 1027
|||||
Db 221 CTGAAGATGTTTTCCTGGGATCAATCTCAGCATTAACACTCTAGTCTAGGTCAAAG 162
|||||
QY 1028 AGGTGATCAAA 1037
|||||
Db 161 AGGTGATTA 152
|||||

RESULT 15
AC102127
LOCUS 196674 bp DNA linear ROD 01-MAY-2004
DEFINITION Mus musculus chromosome 1, clone RP23-202A19, complete sequence.
ACCESSION AC102127.11 GI:46931440
VERSION AC102127.11
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
ChoepeI,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Direct Submission
TITLE
JOURNAL
AUTHORS
```

```
REFERENCE
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
ChoepeI,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hagel,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Submitted (10-APR-2004) Whitehead Institute/MIT Center for Genome
Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
ChoepeI,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hagel,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Submitted (01-MAY-2004) Whitehead Institute/MIT Center for Genome
Direct Submission
TITLE
JOURNAL
COMMENT
On May 1, 2004 this sequence version replaced gi:4638225.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBK
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L18071
Center clone name: 202_A_19
-----
Location/Qualifiers
1. 196674
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP23-202A19"
/clone_lib="RPC1-23 Female Mouse BAC"
-----
FEATURES
source
```

```
misc_feature 1. .4229
/note="wgs_end_extension
clone end:SP6"
repeat_region 1. .193
/rpt_family="MTA"
unsure 614. .746
/note="single clone coverage"
unsure 701. .706
/note="<30 qual single clone coverage"
repeat_region 1208. .1291
/rpt_family="(CTA)n"
repeat_region 1571. .1842
/rpt_family="RLTRETN MM"
repeat_region complement(1974. .2067)
/rpt_family="B4"
repeat_region 3219. .3319
/rpt_family="GA-rich"
repeat_region 3368. .3448
/rpt_family="(CA)n"
repeat_region 3626. .3688
/rpt_family="(CAAA)n"
repeat_region 3925. .4374
/rpt_family="Lx"
misc_feature 4230. .4235
/note="clone boundary
clone end:SP6
site:EcoRI"
repeat_region 4383. .4411
/rpt_family="AT-rich"
repeat_region 4445. .4496
/rpt_family="(TCTCTG)n"
repeat_region 4509. .4638
/rpt_family="(TG)n"
repeat_region 4741. .4784
/rpt_family="(TG)n"
repeat_region 4785. .4832
/rpt_family="(TA)n"
repeat_region 4833. .4955
/rpt_family="(TG)n"
repeat_region 5398. .5431
/rpt_family="(GAAA)n"
repeat_region complement(5434. .5561)
/rpt_family="RMER15"
repeat_region complement(6185. .6272)
/rpt_family="RSINE1"
repeat_region 6516. .6570
/rpt_family="(T)n"
repeat_region complement(6571. .6823)
/rpt_family="Li_MM"
repeat_region 6821. .6939
/rpt_family="Li_MM"
repeat_region 7007. .7285
/rpt_family="Lx5"
repeat_region 7292. .7848
/rpt_family="Lx5"
repeat_region 8187. .8286
/rpt_family="RSINE1"
repeat_region complement(8823. .9061)
/rpt_family="B4A"
repeat_region complement(9243. .9815)
/rpt_family="Li"
repeat_region 9832. .9861
/rpt_family="(A)n"
repeat_region 9862. .10368
/rpt_family="Lx2"
repeat_region complement(10373. .11559)
/rpt_family="Li_MM"
repeat_region 11555. .12196
/rpt_family="Lx2"
repeat_region complement(13221. .13381)
/rpt_family="B3A"
repeat_region 13437. .13578
/rpt_family="B1_MM"
```

```
repeat_region 15222. .15254
/rpt_family="AT rich"
repeat_region complement(15315. .15605)
/rpt_family="RMER20"
repeat_region complement(16008. .16517)
/rpt_family="RMERS"
repeat_region 16673. .16713
/rpt_family="(TG)n"
repeat_region 16809. .16866
/rpt_family="(TG)n"
repeat_region complement(16922. .17052)
/rpt_family="B4"
repeat_region 17101. .17490
/rpt_family="MLT2B1"
repeat_region 17725. .17762
/rpt_family="(GGAA)n"

Query Match 50.2%; Score 522.8; DB 10; Length 196674;
Best Local Similarity 74.4%; Pred. No. 9.5e-129;
Matches 726; Conservative 0; Mismatches 242; Indels 8; Gaps 5;

QY 69 TGTGGGACTGCGGCCATTTGTGGATGGAACAGCGGGGAGCAAGTGATCCCCCTGTCTGGC 128
DB 101632 TTTATGACTATGCCCATTTTATGGATGAAGAAAAAGGAGGATGCCCTTGTGTAAG 101691

QY 129 GGGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGAGACTCTCAA 188
DB 101692 AGGCATATAAAGCTGCTCTTTGAGGTGGCCACCTTGAGGAATCTCTATGAGNATACTAA 101751

QY 189 ATACCAGTTTGTATGCTTCTATGGGAGAGAGTAGTACTAGCAGGAGCATCTATGACAGTTTC 248
DB 101752 ATATCAGTTTAAATGCTTTTATTGAGAGAAGACACTAGCAGAAACATCTAAGTAGGTTTC 101811

QY 249 AGTGAGGCCCATCTTAAGCACTTCTGGAAGGGCAGATGCCAGTGTCTTGCCTATGCG 308
DB 101812 AGTACAGCCTATCTCTGTGSCACTTGTGTGATGGGAGAAATACAGTGTCTTACATATGG 101871

QY 309 ACCCAGAGAGCTGGGAAGACGCACACAATGCTGGGCGACCCC---AGAGCAACCTGGGG 364
DB 101872 GACTATTGAGGCGAGGAGACACATATAATCTTGGGTAGCCCCCTGGGAACACATGGAG 101931

QY 365 TGATCCCGGGCTCTCTATGAGACCTCTGACACTCAAGGAGGAGGAGGTCGCGAGGCC 424
DB 101932 TGATTCCTTAGCCTCTCATAGACCTCTGCAGATAACAAGGGAAGAGTGTGTAAGGCT 101991

QY 425 GGCATGGGCCCTTTCTCTACCATGCTTACTAGAGATCTTACCAGGAGAGGTATTAG 484
DB 101992 GGCATGGGACATTTCTGTAGGATGTCCTATTAGAGGTCTACCAGGAAGAGTATTAG 102051

QY 485 ACCTCTGGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGCCGGGGGAATATCC 544
DB 102052 ACCTCTTATATCTGATCAGTACGAGACCTAGTATATGAGNATGTGCGAGGGAATCC 102111

QY 545 TGATTCGGGTCTCTCCAGAAAGCCCATCAGTAGTCTTGTGCTGATTTTGAAGCGCATTC 604
DB 102112 TGATTCAGTTCTCACCCAGGACCCCAATACTTGTCTCTGACTTTGAGCAGTACTT-C 102170

QY 605 TGCCAGCCAGTCGAATCGGACTGTAGGACCAACCGGCTCAACAGCGCTCTCTCCGCA 664
DB 102171 TGCCAGCCTGTAGAAATCAAGTTGTAGGAGTACTTGTCTTAAGCAGTCTCTTACATA 102230

QY 665 GTCATGCTGTGCTCCTGCTCAAGGTGGACCGGGAACGTTTGGCCCCCATTTTGGCAGC 724
DB 102231 -TCATTCATGCTCTTTATCAAGTTGGATCAGCATGGACATTTGACTCTGTTTCACTAGT 102289

QY 725 GAGAGGAAAACTCTACTGATTGACTTGGTGGGTGAGAGCAACCGCGCACAGGCA 784
DB 102290 GAAAGGAAAACTCTACCATATTGATTGTTGGTTCAGAGGACCAACCAACA-AGGCA 102348

QY 785 ACNAGGCCCTTCGGCTAAAGAGAGTGGGCCATCAACCTCTCCCTGTTTGTCTGGGCA 844
DB 102349 ACCAGTGCAATTCAACTCAAGAGAGTGGGCCATCAACACCT-GTTTCTTGTACTGGGCA 102407
```

```
Qy 845 AAGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGACCTTATCGGACAGCAAGCTCA 904
Db 102408 AAGTGGAGTTGTCATCATATTAGGGCTCCCTCTATATTATCTGGGACAGAAAAATCA 102467
Qy 905 CTCGCCTATTGCAGGACTCTCTGGTGGCTCAGCCACAGTATCCTTATTGCCAACATTG 964
Db 102468 TTTCCTATTGCCAGACTCTCTGGAAGGCTCAGCTCATAACATCATATTACCAACACTG 102527
Qy 965 CCCCTGAGAGACGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCCA 1024
Db 102528 CTCCTGAAGACATTTTACTTGAATCATCTGAGCATTTGAATTTCACTGCTAGGTCCA 102587
Qy 1025 AGGAGGTGATCAATTG 1040
Db 102588 AGGAGCTGATTAATTG 102603
```

Search completed: November 10, 2004, 13:18:05  
Job time : 4604.34 secs



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1041	100.0	1041	10	ADQ23341	Adc23341 DNA encod
2	1041	100.0	1041	12	ADQ60231	Adg60231 Human mic
3	1038.4	99.8	1464	10	ADQ23343	Adc23343 DNA encod
4	1038.4	99.8	1464	12	ADQ60233	Adg60233 Human mic
5	1038	99.7	1115	10	ADQ23337	Adc23337 DNA encod
6	1038	99.7	1115	12	ADQ60227	Adg60227 Human mic
7	1035.4	99.5	1538	10	ADQ23339	Adc23339 DNA encod
8	1035.4	99.5	1538	12	ADQ60229	Adg60229 Human mic
9	1035.4	99.5	2097	12	ADQ09241	Adg09241 Human KNS
10	1035.4	99.5	2099	10	ADK40995	Adk40995 Novel hum
11	1034	99.3	2104	3	AAF15853	Aaf15853 Human pro
12	265	28.5	464	9	ACH44493	Ach44493 Human fce
13	230.2	22.4	531	12	ACH77408	Ach77408 Human gen
14	230.4	22.1	232	12	ACH91108	Ach91108 Human gen
15	184.6	17.7	2095	11	ADMO1564	Admo1564 Human cDN
16	184.6	17.7	3624	12	ADO44181	Ado44181 DNA encod
17	155.8	15.0	3374	10	ADJ95073	Adj95073 Novel NOV
18	154.8	14.9	2375	8	ABA34596	Abx34596 Human mdd
19	136.6	13.1	1152	6	ABA94616	Abx94616 Nucleotid
20	136.6	13.1	1152	8	ABS57216	Abs57216 DNA encod
21	136.6	13.1	4108	6	ABA94614	Abx94614 Human kin

This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardinals, immunomodulators and antiinflammatories. Accordingly, through gene therapy, they can be used for the treatment of cancer, hyperplasias, stenosis, cardiac hypertrophy, immune disorders and inflammation. This polynucleotide sequence is human KID DNA (Seqid 5) encoding a full length KID enzyme of the invention.

```

Q      Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;
      Query Match      100.0%; Score 1041; DB 10; Length 1041;
      Best Local Similarity 100.0%; Pred. No. 5e-296;
      Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

1	QY	ATGGGTCGCTCGCTCGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACTCTCCAGCTCGCGTA	60
1	Db	ATGGGTCGCTCGCTCGGCTAAGCAAGATTGGAGCTACTCGTGTGTCACTCTCCAGCTCGCGTA	60
61	QY	AGGTCGCTGTGCGACTGCGGCCATTTCGTGATGGAAACAGCGGGAGCAAGTCATCCCCCC	120
61	Db	AGGTCGCTGTGCGACTGCGGCCATTTCGTGATGGAAACAGCGGGAGCAAGTCATCCCCCC	120
121	QY	TGTCGGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAAACACACAGGAG	180
121	Db	TGTCGGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAAACACACAGGAG	180
181	QY	ACTCTCAAAATACAGATTGATGCTTCTATGGGAGAGGATCTACGAGGAGCAATCTAT	240
181	Db	ACTCTCAAAATACAGATTGATGCTTCTATGGGAGAGGATCTACGAGGAGCAATCTAT	240
241	QY	GCAGGTTCAAGTCAGCCCATCCTAAGGCATTCCTGGAAGGCGAGAAATGCCAGTGTGTT	300
241	Db	GCAGGTTCAAGTCAGCCCATCCTAAGGCATTCCTGGAAGGCGAGAAATGCCAGTGTGTT	300
301	QY	GCCTATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCAGGCCAGCAACCT	360
301	Db	GCCTATGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCAGGCCAGCAACCT	360
361	QY	GGGTTGATCCCGGGCTCTCATGGAACCTCTCTGACGTCAACAGGAGAGGGTCCGAG	420
361	Db	GGGTTGATCCCGGGCTCTCATGGAACCTCTCTGACGTCAACAGGAGAGGGTCCGAG	420
421	QY	GGCCGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGTCTACAGGAGAGGTA	480
421	Db	GGCCGGCCATGGGCCCTTCTGTCAACCATGCTTACCTAGAGTCTACAGGAGAGGTA	480
481	QY	TTAGACCTCTTGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGATGCGCGGGGAAT	540
481	Db	TTAGACCTCTTGACCCCTGCTTCGGGAGACCTGGTAATCCGAGAAGATGCGCGGGGAAT	540
541	QY	ATCCTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGAGCGGAC	600
541	Db	ATCCTGATTCGGGTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGAGCGGAC	600
601	QY	TTCTCGCAGCCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGGGCTCTCC	660
601	Db	TTCTCGCAGCCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGGGCTCTCC	660
661	QY	CGCAGTCATGTCGTGCTCCTGGTCAAGGTGGACACAGCGGGAACGTTTGCGCCCATTTCCG	720
661	Db	CGCAGTCATGTCGTGCTCCTGGTCAAGGTGGACACAGCGGGAACGTTTGCGCCCATTTCCG	720
721	QY	CAGCGAGGGGAAAATCTCTACTGATTCGCTTGGCTGGGTACAGGACAACCGCGCGACA	780
721	Db	CAGCGAGGGGAAAATCTCTACTGATTCGCTTGGCTGGGTACAGGACAACCGCGCGACA	780

CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents DNA encoding a human microtubule  
 CC motor protein of the invention.  
 XX  
 SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;

Query Match 100.0%; Score 1041; DB 12; Length 1041;  
 Best Local Similarity 100.0%; Pred. No. 5e-296;  
 Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTGCGTGTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCGGTA 60  
 DB 1 ATGGGTGCGTGTGGCTAAGCAAGATTGGAGTACTGCTGCTCCACCTCCAGCTCGGTA 60  
 QY 61 AGGGTGGCTGTGGAGTCTGGGCGCATTTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120  
 DB 61 AGGGTGGCTGTGGAGTCTGGGCGCATTTGTGGATGGAACAGCGGGAGCAAGTATCCCCC 120  
 QY 121 TGTGTGGGGGCGATGAGACAGTCTCTCTAGAGATTGCTTAATGAGGAACACACAGGAG 180  
 DB 121 TGTGTGGGGGCGATGAGACAGTCTCTCTAGAGATTGCTTAATGAGGAACACACAGGAG 180  
 QY 181 ACTCTCAATACAGTTGATGCTCTTATGSGGAGAGGATCTACAGAGGACATCTAT 240  
 DB 181 ACTCTCAATACAGTTGATGCTCTTATGSGGAGAGGATCTACAGAGGACATCTAT 240  
 QY 241 GCAGGTTTCAGTCAGCGCCCATCTTAAGGCATCTCTGGAAGGGCAGAGTGCAGTGTCTT 300  
 DB 241 GCAGGTTTCAGTCAGCGCCCATCTTAAGGCATCTCTGGAAGGGCAGAGTGCAGTGTCTT 300  
 QY 301 GCCTATGAGCCCAACAGAGTGGGAAGAGCACAATGCTGGGCGAGCCCGAGCAACCT 360  
 DB 301 GCCTATGAGCCCAACAGAGTGGGAAGAGCACAATGCTGGGCGAGCCCGAGCAACCT 360  
 QY 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGTGCCGAG 420  
 DB 361 GGGGTGATCCCGGGGCTCTCATGGACCTCTCTGAGCTCACAAGGAGGAGGTGCCGAG 420  
 QY 421 GCGCGGCCATGGGCCCTTTCTGTACCATGCTTTACTAGAGATCTACAGGAGAGGTA 480  
 DB 421 GCGCGGCCATGGGCCCTTTCTGTACCATGCTTTACTAGAGATCTACAGGAGAGGTA 480  
 QY 481 TTAGACTCTGGAACCTGCTTCCGAGAGCCTGGTAATCCGAGAAGACTGCCGGGGAAT 540  
 DB 481 TTAGACTCTGGAACCTGCTTCCGAGAGCCTGGTAATCCGAGAAGACTGCCGGGGAAT 540  
 QY 541 ATCCTGATTCGGGCTCTCCAGAGGCCATCAGTAGCTTGTGATTTGAGCGGCAC 600  
 DB 541 ATCCTGATTCGGGCTCTCCAGAGGCCATCAGTAGCTTGTGATTTGAGCGGCAC 600  
 QY 601 TTCCTCCAGCCAGTCGAATCGGACTGTAGAGGCCACCGGCTCAACAGCGCTCCTCC 660  
 DB 601 TTCCTCCAGCCAGTCGAATCGGACTGTAGAGGCCACCGGCTCAACAGCGCTCCTCC 660  
 QY 661 CGCAGTCATGCTGTGCTCTGCTGCAAGGTGGACAGCGGGAACGTTTGGCCCCATTTCCG 720  
 DB 661 CGCAGTCATGCTGTGCTCTGCTGCAAGGTGGACAGCGGGAACGTTTGGCCCCATTTCCG 720  
 QY 721 CAGCGAGAGGGAACCTTACTGATTTGATCTGGGTGAGGACCAACCGGCGCACA 780  
 DB 721 CAGCGAGAGGGAACCTTACTGATTTGATCTGGGTGAGGACCAACCGGCGCACA 780

QY 781 GGCAACAAGGCGCTTGGCTAAAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 840  
 DB 781 GGCAACAAGGCGCTTGGCTAAAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTG 840  
 QY 841 GGCAAGTGTAGATGCGCTGAATCAAGGGCTCCCTGCTGTACTTATCGGGACAGCAAG 900  
 DB 841 GGCAAGTGTAGATGCGCTGAATCAAGGGCTCCCTGCTGTACTTATCGGGACAGCAAG 900  
 QY 901 CTCACCTCGCTATTGAGGAGTCTCTGGGTGGCTCAGCCACAGTATCCTTATTTGCCAAC 960  
 DB 901 CTCACCTCGCTATTGAGGAGTCTCTGGGTGGCTCAGCCACAGTATCCTTATTTGCCAAC 960  
 QY 961 ATTGCCCTCGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020  
 DB 961 ATTGCCCTCGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020  
 QY 1021 TCCAAGGAGTGAATCAATTGA 1041  
 DB 1021 TCCAAGGAGTGAATCAATTGA 1041

RESULT 3  
 ADC23343

ID ADC23343 standard; DNA; 1464 BP.

XX ADC23343;

AC AC (first entry)

DT 18-DEC-2003 (first entry)

XX DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).  
 XX human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
 XX cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
 XX cancer; hyperplasia; restenosis; cellular proliferation disorder;  
 XX cardiac hypertrophy; immune disorder; inflammation.

OS Homo sapiens.

Key Location/Qualifiers  
 PH 1.1464  
 FT CDS

FT /\*tag= a  
 FT /product= "KID protein"

XX US6387644-B1.

XX 14-MAY-2002.

XX 28-NOV-2000; 2000US-00724224.

XX 20-APR-1999; 99US-00295612.

XX 20-JUN-2000; 2000US-00597292.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C;

XX WPI; 2003-706919/67.

XX P-PSDB; ADC23344.

XX Identifying a candidate agent as modulator of function of a target  
 XX protein for treating cellular proliferation disorders by adding a  
 XX candidate agent to a mixture of the target protein that  
 XX directly/indirectly produces ADP or phosphate.

XX Disclosure; SEQ ID NO 7; 26pp; English.

XX This invention relates to a novel method for high throughput screening  
 XX systems used to identify compounds for the treatment of cellular  
 XX proliferation disorders. Specifically, it refers to candidate agents that  
 XX are capable of modulating the activity of target proteins having motor  
 XX domains, such that the target protein directly or indirectly produces ADP  
 XX or phosphate. Furthermore, this activity can be determined using



screening for modulators of motor proteins useful for treating cellular proliferation disorders such as cancer, hyperplasia, restenosis, cardiac hypertrophy, immune disorders, and inflammation for treating disorders associated with KIP and for inhibiting KIP for treating autoimmune diseases, arthritis, graft rejection, inflammatory disorders, and proliferation induced after medical procedures including surgery and angioplasty. This sequence represents DNA encoding a human microtubule motor protein of the invention.

```
Query Match          99.8%; Score 1038.4; DB 12; Length 1464;
Best Local Similarity 99.9%; Pred. No. 3.4e-295;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	1	ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGTACTCGTGTGCACCTCCAGCTCGCGTA	60
Db	1	ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGTACTCGTGTGCACCTCCAGCTCGCGTA	60
QY	61	AGGTTGGCTGTGCGACTGCGCCATTTCGTGATGGAAACAGCGGGAGCAAGTGATCCCC	120
Db	61	AGGTTGGCTGTGCGACTGCGCCATTTCGTGATGGAAACAGCGGGAGCAAGTGATCCCC	120
QY	121	TGTTGCGGGCATGGAACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACACAGGAG	180
Db	121	TGTTGCGGGCATGGAACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAACACAGGAG	180
QY	181	ACTCTCAATATACAGTTTGTATGCTTTATGGGAGAGGATCTACAGAGGACATCTAT	240
Db	181	ACTCTCAATATACAGTTTGTATGCTTTATGGGAGAGGATCTACAGAGGACATCTAT	240
QY	241	GCAGTTTCAGTGACGCCCATCTTAAGGCACTTGCTTGGAAAGGCGAGAAATGCCAGTGTGCTT	300
Db	241	GCAGTTTCAGTGACGCCCATCTTAAGGCACTTGCTTGGAAAGGCGAGAAATGCCAGTGTGCTT	300
QY	301	GCCTATGGACCCACAGGAGCTGGGAAGACGCACAAATGCTGGGCAGCCACAGAGCAACCT	360
Db	301	GCCTATGGACCCACAGGAGCTGGGAAGACGCACAAATGCTGGGCAGCCACAGAGCAACCT	360
QY	361	GGGTTGATCCGCGGGCTCTCATGGACCTCTGTGACGTCAACAGGAGGAGGGTGCCGAG	420
Db	361	GGGTTGATCCGCGGGCTCTCATGGACCTCTGTGACGTCAACAGGAGGAGGGTGCCGAG	420
QY	421	GGCGGGCATGGGGCCCTTTCGTGCACCAATGCTTACCTAGAGATCTACACAGAGGAAGTA	480
Db	421	GGCGGGCATGGGGCCCTTTCGTGCACCAATGCTTACCTAGAGATCTACACAGAGGAAGTA	480
QY	481	TTAGACTCTCTGACCTCTGCTTCGGGAGACTGTGTAATCCGAGAAGACTGCGGGGGAAAT	540
Db	481	TTAGACTCTCTGACCTCTGCTTCGGGAGACTGTGTAATCCGAGAAGACTGCGGGGGAAAT	540
QY	541	ATCTTGATTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGGGGCAC	600
Db	541	ATCTTGATTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGGGGCAC	600
QY	601	TTCTGTCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACGCGTCTCTCC	660
Db	601	TTCTGTCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACGCGTCTCTCC	660
QY	661	CGCAGTCATGCTGTGCTCTCTGGTCAAGGTGGACACAGCGGGAACGTTTGGCCCCATTTGCG	720
Db	661	CGCAGTCATGCTGTGCTCTCTGGTCAAGGTGGACACAGCGGGAACGTTTGGCCCCATTTGCG	720
QY	721	CACGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTACAGAGGACACACCGGGGCACA	780
Db	721	CACGAGAGGGAACCTCTACCTGATTGACTTGGGTGGGTACAGAGGACACACCGGGGCACA	780
QY	781	GGCAACAAGGGCCCTTCGGCTAAAAGAGAGTGGAGCCATCAACACTCCCTGTTGTCTGTG	840
Db	781	GGCAACAAGGGCCCTTCGGCTAAAAGAGAGTGGAGCCATCAACACTCCCTGTTGTCTGTG	840
QY	841	GGCAAAAGTGTGTAGTGGCTGAATCAGGGGCTCCCTCGTGTACTTATCGGGACAGCAAG	900

Db	841	GGCAAGTGTAGATGGCTGAATCAGGGGCTCCCTCGTGTA	CCTTATCGGACAGCAAG	900
Qy	901	CTCACTGGCCTATTGACGAGACTCTCTGGGTGGCTCAGCCACACAGTAT	TCCTTATTGGCAAC	960
Db	901	CTCACTGGCCTATTGACGAGACTCTCTGGGTGGCTCAGCCACACAGTAT	TCCTTATTGGCAAC	960
Qy	961	ATTGCCCTCAGAGAGCGCTTCTACCTAGACACACAGTCTCCGCACT	CAAACTTTGCTGCCAGG	1020
Db	961	ATTGCCCTCAGAGAGCGCTTCTACCTAGACACACAGTCTCCGCACT	CAAACTTTGCTGCCAGG	1020
Qy	1021	TCCAAGGAGGTGATCAATTG	1040	
Db	1021	TCCAAGGAGGTGATCAATCG	1040	
RESULT 5				
ADC23337				
ID	ADC23337 standard; DNA: 1115 BP.			

US6387644-B1.

14-MAY-2002

28-NOV-2000: 2000US-00724224.

20-APR-1999: 99US-00295612.

20-JUN-2000; 2000US-00597292.

(CYTO-) CYTOKINETICS INC.

Beraud C;

WPI; 2003-706919/67.

[illegible]

identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.

Disclosure; SEQ ID NO 1; 26pp; English.

This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cycostatic, cardiants, immunomodulators and antiinflammatoryes. Accordingly, through gene

CC therapy, they can be used for the treatment of cancer, hyperplasias,  
 CC restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
 CC polynucleotide sequence is human KID DNA (Seqid 1) encoding a full length  
 CC KID enzyme of the invention.  
 XX

SQ Sequence 1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;  
 Query Match 99.7%; Score 1038; DB 10; Length 1115;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-295;  
 Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGTGCTGTCGCTAAGCAAGATTGGAGCTACTGTCGTCACCTCCAGCTCGCGTAAGG 63  
 DB 78 GGTGCTGTCGCTAAGCAAGATTGGAGCTACTGTCGTCACCTCCAGCTCGCGTAAGG 137  
 QY 64 GTGGCTGTCGCTAAGCAAGATTGGAGCTACTGTCGTCACCTCCAGCTCGCGTAAGG 123  
 DB 138 GTGGCTGTCGCTAAGCAAGATTGGAGCTACTGTCGTCACCTCCAGCTCGCGTAAGG 197  
 QY 124 GTGGCTGTCGCTAAGCAAGATTGGAGCTACTGTCGTCACCTCCAGCTCGCGTAAGG 183  
 DB 198 GTGGCTGTCGCTAAGCAAGATTGGAGCTACTGTCGTCACCTCCAGCTCGCGTAAGG 257  
 QY 184 CTCAATACCACTTGTATGCTCTCTATGGGAGGAGTACTCAGCAGCATCTATGCA 243  
 DB 258 CTCAATACCACTTGTATGCTCTCTATGGGAGGAGTACTCAGCAGCATCTATGCA 317  
 QY 244 GGTTCAGTGCAGCCCATCTAAGCACTTGTGGAAGGCGAGATGCGAGTGTCTGCC 303  
 DB 318 GGTTCAGTGCAGCCCATCTAAGCACTTGTGGAAGGCGAGATGCGAGTGTCTGCC 377  
 QY 304 TATGACCCACAGAGCTGGAGAGCAGACACATGCTGGGAGGCGGAGCAACCTGGG 363  
 DB 378 TATGACCCACAGAGCTGGAGAGCAGACACATGCTGGGAGGCGGAGCAACCTGGG 437  
 QY 364 GTGATCCCGCGGCTCTCATGACCTCTGTCAGCTCACAAGGGAGGAGGTCGCGAGGCG 423  
 DB 438 GTGATCCCGCGGCTCTCATGACCTCTGTCAGCTCACAAGGGAGGAGGTCGCGAGGCG 497  
 QY 424 CGGCCATGGGCTTCTGTACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 483  
 DB 498 CGGCCATGGGCTTCTGTACCATGCTTACCTAGAGATCTACAGAGAGAGTATTA 557  
 QY 484 GACCTCTGGACCTCTTCGGAGACCTGTAATCCGAGAGAGCTGCGGGGAGTATC 543  
 DB 558 GACCTCTGGACCTCTTCGGAGACCTGTAATCCGAGAGAGCTGCGGGGAGTATC 617  
 QY 544 CTGATTCGGGCTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTC 603  
 DB 618 CTGATTCGGGCTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTC 677  
 QY 604 CTGCGAGCAGTCGAATCGATCTGAGAGCCACCGGCTCAACAGGGCTCTCCGCG 663  
 DB 678 CTGCGAGCAGTCGAATCGATCTGAGAGCCACCGGCTCAACAGGGCTCTCCGCG 737  
 QY 664 AGTCATGCTGCTCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 723  
 DB 738 AGTCATGCTGCTCTGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797  
 QY 724 CGAGAGGAGAACTCTACCTGATTTGATTTGGTGGGTGAGAGCAACCGGGGCGACAGG 783  
 DB 798 CGAGAGGAGAACTCTACCTGATTTGATTTGGTGGGTGAGAGCAACCGGGGCGACAGG 857  
 QY 784 AACAGGGCTTGGCTAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 843  
 DB 858 AACAGGGCTTGGCTAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 917  
 QY 844 AAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903  
 DB 918 AAGTGTAGATCGCTGAATCAGGGCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 977  
 QY 904 ACTCGCTATTGAGAGCTCTCTGGGTGGGTGAGCCACAGTATCTTATTCGCAACATT 963

DB 978 ACTCGCTATTGCGAGCTCTCTGGGTGGTCCAGCCACAGTATCTTATTGCCAACATT 1037  
 QY 964 GCCCTTGAGAGAGCTTCTACTAGACACAGTCTCGGACTCACTTCTGTCGAGGTCC 1023  
 DB 1038 GCCCTTGAGAGAGCTTCTACTAGACACAGTCTCGGACTCACTTCTGTCGAGGTCC 1097  
 QY 1024 AAGGAGGTGATCAATTGA 1041  
 DB 1098 AAGGAGGTGATCAATTGA 1115  
 RESULT 6  
 ADQ60227  
 ID ADQ60227 standard; DNA; 1115 BP.  
 XX AC ADQ60227;  
 XX AC  
 DT 23-SEP-2004 (first entry)  
 XX Human microtubule motor protein DNA #1.  
 DE Human; microtubule motor protein; gene; ds;  
 KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
 KW cardiac hypertrophy; immune disorder; inflammation;  
 KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;  
 KW graft rejection; inflammatory bowel disease; angioplasty.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US6762043-B1.  
 XX 13-JUL-2004.  
 XX  
 XX 06-MAR-2002; 2002US-00093317.  
 XX  
 XX 20-APR-1999; 99US-002995612.  
 XX 20-JUN-2000; 2000US-00597292.  
 XX 28-NOV-2000; 2000US-00724224.  
 XX  
 XX (CYTO-) CYTOKINETICS INC.  
 XX  
 XX Beraud C;  
 XX  
 XX WPI; 2004-532491/51.  
 DR  
 XX New isolated microtubule motor protein, useful for screening modulators  
 PT for treating cellular proliferation disorders such as cancer,  
 PT hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 PT inflammation.  
 XX  
 PS Disclosure; SEQ ID NO 1; 26pp; English.  
 CC  
 CC The invention relates to human microtubule motor proteins and the nucleic  
 CC acids encoding them. The invention also relates to a method of screening  
 CC for modulators of a motor protein which has microtubule stimulated ATPase  
 CC activity, a method of testing for ATPase activity of microtubule motor  
 CC proteins, methods to identify candidate agents that bind to a target  
 CC protein or act as a modulator of the binding characteristics or  
 CC biological activity of a target protein, modulators of the target  
 CC protein, and methods of treating cellular proliferation disorders such as  
 CC cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 CC and inflammation, for treating disorders associated with kinesin-like DNA  
 CC binding protein (KID) and for inhibiting KID. The sequences are used for  
 CC screening for modulators of motor proteins useful for treating cellular  
 CC proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 CC hypertrophy, immune disorders and inflammation, for treating disorders  
 CC associated with KID and for inhibiting KID and for treating autoimmune  
 CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
 CC proliferation induced after medical procedures including surgery and  
 CC angioplasty. This sequence represents DNA encoding a human microtubule  
 CC motor protein of the invention. Note: The specification states that this  
 CC sequence encodes the protein featured as SEQ ID NO:2, but this does not  
 CC appear to be the case.

XX	Sequence	1115 BP; 241 A; 318 C; 333 G; 223 T; 0 U; 0 Other;	
QQ	Query Match	99.7%; Score 1038; DB 12; Length 1115;	
QQ	Best Local Similarity	100.0%; Pred. No. 3.9e-295;	
QQ	Matches 1038; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	4	GTCGCTGTGCGTAAAGATGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG	63
DB	78	GGTCGCTGTGCGTAAAGATGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG	137
QY	64	GTGCTGTGCGTAAAGATGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG	123
DB	138	GTGCTGTGCGTAAAGATGGAGTACTCGTCCACCTCCAGCTCGCGTAAGG	197
QY	124	GTGCGGCGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGCT	183
DB	198	GTGCGGCGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGCT	257
QY	184	CTCAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
DB	258	CTCAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
QY	244	GTTTCAGTCAGCCCATCTTAAGCAGCTTCTGGAAGGAGGAGATGCGAGTGTGCTGCC	303
DB	318	GTTTCAGTCAGCCCATCTTAAGCAGCTTCTGGAAGGAGGAGATGCGAGTGTGCTGCC	377
QY	304	TATGGACCCACAGGAGCTGGAGAGCAGACACAACTGCTGGGAGGAGGAGGAGGAGG	363
DB	378	TATGGACCCACAGGAGCTGGAGAGCAGACACAACTGCTGGGAGGAGGAGGAGGAGG	437
QY	364	GTGATCCCGCGGCTCTCATGGACCTCTGAGCTCTCAGCTCAAGAGGAGGAGGAGGAGG	423
DB	438	GTGATCCCGCGGCTCTCATGGACCTCTGAGCTCTCAGCTCAAGAGGAGGAGGAGGAGG	497
QY	424	CGGCCATGGGCGCTTTCTGTCACCATGCTCTTACCTAGAGATCTACAGGAGAGGATTA	483
DB	498	CGGCCATGGGCGCTTTCTGTCACCATGCTCTTACCTAGAGATCTACAGGAGAGGATTA	557
QY	484	GACCTCTGAGCCTGCTTCCGAGAGCTGCTGATTCGAGAGAGTCTCAGGAGGAGATATC	543
DB	558	GACCTCTGAGCCTGCTTCCGAGAGCTGCTGATTCGAGAGAGTCTCAGGAGGAGATATC	617
QY	544	CTGATTCGCGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC	603
DB	618	CTGATTCGCGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC	677
QY	604	CTGCCAGCAGTGAATCGGAGCTGTAGAGCCACCGGCTCAACAGCGCTCTCCCGC	663
DB	678	CTGCCAGCAGTGAATCGGAGCTGTAGAGCCACCGGCTCAACAGCGCTCTCCCGC	737
QY	664	AGTCATGCTGCTCTGCTCAAGGTGGACCGGAGGAGTGTGGCCCATTTGCCAG	723
DB	738	AGTCATGCTGCTCTGCTCAAGGTGGACCGGAGGAGTGTGGCCCATTTGCCAG	797
QY	724	CGAGAGGAGAACTCTACCTGATGACTTGGCTGGGTGAGAGCAACCGGCGCAGAGC	783
DB	798	CGAGAGGAGAACTCTACCTGATGACTTGGCTGGGTGAGAGCAACCGGCGCAGAGC	857
QY	784	AACAAGGCGCTTCGGCTAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC	843
DB	858	AACAAGGCGCTTCGGCTAAAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC	917
QY	844	AAAGTGTAGATCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGAGCAGAGCTC	903
DB	918	AAAGTGTAGATCGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGAGCAGAGCTC	977
QY	904	ACTCGCTATTGCGAGCTCTCTGGGTGGCTCAGCCACAGATATCTTATGCGCAACT	963
DB	978	ACTCGCTATTGCGAGCTCTCTGGGTGGCTCAGCCACAGATATCTTATGCGCAACT	1037
QY	964	GCCCTGTAGAGCGCTTCTACCTTAGACAGAGTCTCGGACTCAACTTGTGCGCAGGTC	1023

DB	1038	GCCCTCGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTCC	1097
QY	1024	AAGGAGGTGATCAATTGA	1041
DB	1098	AAGGAGGTGATCAATTGA	1115
RESULT 7			
ADC23339			
ID	ADC23339	standard; DNA; 1538 BP.	
XX	ADC23339;		
AC	ADC23339;		
XX	18-DEC-2003	(first entry)	
DT			
XX	DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 3).		
DE			
XX	human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;		
KW	cytostatic; cardiac; immunomodulator; antiinflammatory; gene therapy;		
KW	cancer; hyperplasia; restenosis; cellular proliferation disorder;		
KW	cardiac hypertrophy; immune disorder; inflammation.		
XX	Homo sapiens.		
OS			
XX			
Key	Location/Qualifiers		
CDS	1..1538		
FT	/*tag= a		
FT	/product= "KID protein"		
FT	/transl_except= (pos: 4..5; aa: Pro)		
FT	/note= "This codon has an apparent 1 nucleotide deletion		
FT	that alters the reading frame"		
XX	US637644-B1.		
XX	14-MAY-2002.		
XX	28-NOV-2000; 2000US-00724224.		
XX	20-APR-1999; 99US-00295612.		
PR	20-JUN-2000; 2000US-00597292.		
XX	(CYTO-) CYTOKINETICS INC.		
PA			
XX	Beraud C;		
PI	WPI; 2003-706919/67.		
XX	P-PSDB; ADC23340.		
DR			
XX	Identifying a candidate agent as modulator of function of a target		
PT	protein for treating cellular proliferation disorders by adding a		
PT	candidate agent to a mixture of the target protein that		
PT	directly/indirectly produces ADP or phosphate.		
XX	Disclosure; SEQ ID NO 3; 26pp; English.		
XX			
CC	This invention relates to a novel method for high throughput screening		
CC	systems used to identify compounds for the treatment of cellular		
CC	proliferation disorders. Specifically, it refers to candidate agents that		
CC	are capable of modulating the activity of target proteins having motor		
CC	domains, such that the target protein directly or indirectly produces ADP		
CC	or phosphate. Furthermore, this activity can be determined using		
CC	fluorescence or absorbance readouts. The present invention describes a		
CC	method that identifies modulators of the target protein, which is a		
CC	kinesin-like DNA binding protein (known as KID) as cytostatic, cardiac,		
CC	immunomodulators and antiinflammatories. Accordingly, through gene		
CC	therapy, they can be used for the treatment of cancer, hyperplasia,		
CC	restenosis, cardiac hypertrophy, immune disorders and inflammation. This		
CC	polynucleotide sequence is human KID DNA (SeqID 3) encoding a full length		
CC	KID enzyme of the invention.		
XX	Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;		
QQ	Query Match	99.5%; Score 1035.4; DB 10; Length 1538;	



QY	4	GGTTCGCTGTGGCTAAGCAAGATTGGAGCTTACTCGTCGTCACCTCCAGCTCGCGTAAGG	63
Db	78	GGTTCGCTGTGGCTAAGCAAGATTGGAGCTTACTCGTCGTCACCTCCAGCTCGCGTAAGG	137
QY	64	GTGGCTGTGGCACTGGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGATGATCCCCCTGT	123
Db	138	GTGGCTGTGGCACTGGCGCCATTGTGGATGGAAACAGCGGAGCAAGTGATGATCCCCCTGT	197
QY	124	GTGCGGGGCATGACACAGCTGCTCTTAGAGATTGCTTAACTGGAGGAACACACAGGAGACT	183
Db	198	GTGCGGGGCATGACACAGCTGCTCTTAGAGATTGCTTAACTGGAGGAACACACAGGAGACT	257
QY	184	CTCAAAATACCAGTTTGATGCTCTTATGGGGAGAGGACTACTCAGCAGGACATCTTATGCA	243
Db	258	CTCAAAATACCAGTTTGATGCTCTTATGGGGAGAGGACTACTCAGCAGGACATCTTATGCA	317
QY	244	GGTTCAGTGAGGCCCATCTTAAGGCACTTGCTGGAGGGCAGAAATGCCAGTGTCGTGCC	303
Db	318	GGTTCAGTGAGGCCCATCTTAAGGCACTTGCTGGAGGGCAGAAATGCCAGTGTCGTGCC	377
QY	304	TATGGACCCACAGAGCTGGGAAGACGCAACAATGCTGGCGACGCCAGAGCAACCTGGG	363
Db	378	TATGGACCCACAGAGCTGGGAAGACGCAACAATGCTGGCGACGCCAGAGCAACCTGGG	437
QY	364	GTGATCCCGCGGGCTCTCATGAGCACTCTCGCAGTCTCAAGAGGAGAGGGTGCAGGGCC	423
Db	438	GTGATCCCGCGGGCTCTCATGAGCACTCTCGCAGTCTCAAGAGGAGAGGGTGCAGGGCC	497
QY	424	CGGCCATGGGCCCTTTCTGTCACCATGCTCTTACTAGAGATCTTACACAGGAGAAAGGTATTA	483
Db	498	CGGCCATGGGCCCTTTCTGTCACCATGCTCTTACTAGAGATCTTACACAGGAGAAAGGTATTA	557
QY	484	GACCTCTTGACCCCTGCTTCGGGAGACCTGGTAAATCGAGAAAGCTGCGGGGGAATATC	543
Db	558	GACCTCTTGACCCCTGCTTCGGGAGACCTGGTAAATCGAGAAAGCTGCGGGGGAATATC	617
QY	544	CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTCGTGATTTTGAGCGGCACCTC	603
Db	618	CTGATTCGGGGTCTCTCCAGAGACCCATCAGTAGCTTTCGTGATTTTGAGCGGCACCTC	677
QY	604	CTGCCAGCCAGTCGAATTCGCACTGTAGGAGCCACCCGGCTCAACACAGCGCTCCTCCCGC	663
Db	678	CTGCCAGCCAGTCGAATTCGCACTGTAGGAGCCACCCGGCTCAACACAGCGCTCCTCCCGC	737
QY	664	AGTCATGCTGTGCTCTGCTCAAGGTGGACCGAGGGGAAAGTTTGGCCCCCATTTCCGCAG	723
Db	738	AGTCATGCTGTGCTCTGCTCAAGGTGGACCGAGGGGAAAGTTTGGCCCCCATTTCCGCAG	797
QY	724	CGAGAGGGAACCTCTACCTGATTGACTTTGGCTGGGTCTAGAGGACAAACCCGCGCACAGGC	783
Db	798	CGAGAGGGAACCTCTACCTGATTGACTTTGGCTGGGTCTAGAGGACAAACCCGCGCACAGGC	857
QY	784	AACAAGGGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGCC	843
Db	858	AACAAGGGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGCC	917
QY	844	AAAGTGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGATCCTTATCGGGACAGCAAGCTC	903
Db	918	AAAGTGTGTAGATGCGCTGAATCAGGGCTCCCTCGTGATCCTTATCGGGACAGCAAGCTC	977
QY	904	ACTCGCCTATTGACAGGACTCTCTGGGTGGCTCAGGCCACAGTAGTCTTTATGGCCAACTT	963
Db	978	ACTCGCCTATTGACAGGACTCTCTGGGTGGCTCAGGCCACAGTAGTCTTTATGGCCAACTT	1037
QY	964	GCCCTGTAGACAGCTTCTACTAGACACAGTCTCCGACCTCAACTTGTGTCGACAGTCC	1023
Db	1038	GCCCTGTAGACAGCTTCTACTAGACACAGTCTCCGACCTCAACTTGTGTCGACAGTCC	1097
QY	1024	AAGGAGGTGATCAATTG	1040



Db 78 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCTAAGG 137  
Qy 64 GTGGCTGTGCGACTCGCGCCGATTTGTGGATGGAAACAGCGGAGCAAGTGAATCCCTCTGT 123  
Db 138 GTGGCTGTGCGACTCGCGCCGATTTGTGGATGGAAACAGCGGAGCAAGTGAATCCCTCTGT 197  
Qy 124 GTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 183  
Db 198 GTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 257  
Qy 184 CTCAAATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGACGACATCTATGCA 243  
Db 258 CTCAAATACAGTTTGTATGCTCTTATGGGAGAGAGTACTCAGACGACATCTATGCA 317  
Qy 244 GGTTCAGTGCAGCCCATCTTAAGGCACTTCTGTAAGGGGAGAGTGCACGTGCTTGGC 303  
Db 318 GGTTCAGTGCAGCCCATCTTAAGGCACTTCTGTAAGGGGAGAGTGCACGTGCTTGGC 377  
Qy 304 TATGGAACCCACAGGAGCTGGAAAGACGACACAATGCTGGGAGCCGACAGCAACCTGGG 363  
Db 378 TATGGAACCCACAGGAGCTGGAAAGACGACACAATGCTGGGAGCCGACAGCAACCTGGG 437  
Qy 364 GTGATCCCGGGGCTCTCATGACCTCTCTGACCTCACAAGGAGGAGGTGCGGAGGC 423  
Db 438 GTGATCCCGGGGCTCTCATGACCTCTCTGACCTCACAAGGAGGAGGTGCGGAGGC 497  
Qy 424 CGGCCATGGGCCCTTTCTGTACCATGCTCTTACTAGAGATCTACCAAGAGAGGTATTA 483  
Db 498 CGGCCATGGGCCCTTTCTGTACCATGCTCTTACTAGAGATCTACCAAGAGAGGTATTA 557  
Qy 484 GACCTCTGAGACCTCTGCTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGAAATATC 543  
Db 558 GACCTCTGAGACCTCTGCTGGGAGACCTGGTAAATCCGAGAGACTGCCGGGGAAATATC 617  
Qy 544 CTGATTCGGGCTCTCCGAAAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 603  
Db 618 CTGATTCGGGCTCTCCGAAAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 677  
Qy 604 CTGCGAGGAGTGAATCGGACTGTAGGAGCCACCGGCTCAACAGGAGCTCTCCCGC 663  
Db 678 CTGCGAGGAGTGAATCGGACTGTAGGAGCCACCGGCTCAACAGGAGCTCTCCCGC 737  
Qy 664 AGTCATGCTGTCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCAATTTGCCAG 723  
Db 738 AGTCATGCTGTCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCAATTTGCCAG 797  
Qy 724 CGAGAGGGAATCTTACCTGATTTGACTTGGCTGAGAGCAACCGGGGCAAGGC 783  
Db 798 CGAGAGGGAATCTTACCTGATTTGACTTGGCTGAGAGCAACCGGGGCAAGGC 857  
Qy 784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTGGGC 843  
Db 858 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTGGGC 917  
Qy 844 AAAGTGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 903  
Db 918 AAAGTGTAGATCGCTGAATCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTC 977  
Qy 904 ACTGCTTATGAGGACTCTCTGGGTGCTCAGGCCACAGTATCTTATGCGCAACTT 963  
Db 978 ACTGCTTATGAGGACTCTCTGGGTGCTCAGGCCACAGTATCTTATGCGCAACTT 1037  
Qy 964 GCGCTGAGAGAGCTTCTACCTAGACAGCTCTCGGACTCAACTTCTGCTGCGAGTCC 1023  
Db 1038 GCGCTGAGAGAGCTTCTACCTAGACAGCTCTCGGACTCAACTTCTGCTGCGAGTCC 1097  
Qy 1024 AAGGAGTGCATCAATTG 1040  
Db 1098 AAGGAGTGCATCAATCG 1114

ADQ09241

ID ADQ09241 standard; cDNA; 2097 BP.

XX AC ADQ09241;

XX DT 23-SBP-2004 (first entry)

XX XX Human KNSL4 encoding cDNA SEQ ID NO:426.

XX DE thanatos-associated protein; THAP; THAP responsive gene; THAP family;  
XX KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;  
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;  
KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;  
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;  
KW human; gene; ss.

XX OS Homo sapiens.

XX PN WO2004055050-A2.

XX PD 01-JUL-2004.

XX PF 10-DEC-2003; 2003WO-IB006434.

XX PR 10-DEC-2002; 2002US-0432699P.

XX PR 03-JUL-2003; 2003US-0485027P.

XX PA (ENDO-) ENDOCTUBE SAS.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX PI Girard J, Amalric F, Roussigne M, Clouaire T;

XX DR WPI: 2004-525034/50.

XX DR P-PSDB; ADQ09240.

XX PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)  
PT responsive gene for preventing or treating e.g. cancer or inflammation,  
PT comprises modulating the interaction of a THAP polypeptide with a nucleic  
PT acid.

XX PS Example 47; SEQ ID NO 426; 612pp; English.

XX CC The present invention describes a method for modulating the expression of  
CC a thanatos (death)-associated protein (THAP) responsive gene. The method  
CC comprises modulating the interaction of a THAP-family polypeptide or its  
CC biological fragment with a nucleic acid, and so enhancing or repressing  
CC the expression of the THAP responsive gene. Also described: (1) a method  
CC of modulating the expression of a gene responsive to a THAP/chemokine  
CC complex; (2) a pharmaceutical composition comprising a THAP responsive  
CC element in a pharmaceutical carrier; (3) a transcription factor decoy  
CC consisting essentially of a THAP responsive element; (4) a cell  
CC comprising a transcription factor decoy described above; (5) methods of  
CC modulating the interaction between a nucleic acid and a THAP-family  
CC polypeptide or its biological fragment, or a nucleic acid and a  
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a  
CC cell comprising a viral vector which comprises a promoter operably linked  
CC to a nucleic acid encoding a THAP-family polypeptide or its biological  
CC fragment; (7) a method of constructing a cell which expresses a  
CC recombinant THAP-family polypeptide; (8) a method of ameliorating  
CC symptoms associated with a condition mediated by a THAP/chemokine complex  
CC; (9) methods of identifying a test compound that modulates transcription  
CC of a THAP responsive element or that modulates the transport of a  
CC chemokine into the nucleus; (10) methods for reducing the symptoms  
CC associated with a condition selected from excessive or insufficient  
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive  
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative  
CC diseases; symptoms associated with a condition resulting from the  
CC activity of a chemokine or a THAP-family polypeptide in an individual; or  
CC symptoms associated with transcriptional repression or activation  
CC mediated by a THAP-family polypeptide in an individual; (11) a vector  
CC comprising a THAP responsive promoter operably linked to a nucleic acid  
CC encoding a detectable product; (12) a genetically engineered cell  
CC comprising the vector described above or that expresses a THAP-family

polypeptide or its biological fragment; (13) an in vitro transcription reaction comprising a nucleic acid comprising a THAP responsive promoter, ribonucleotides and an RNA polymerase; and (14) an isolated human THAP family polypeptide that does not bind to a chemokine. The pharmaceutical composition has antiangiogenic, antiinflammatory, cardiovascular, cytoskeletal, neuroprotective and osteopathic activities, and can be used as a THAP and THAP synthesis modulator. The composition can be used for modulating the expression of a THAP responsive gene. Modulation is useful for reducing symptoms of conditions such as excessive or insufficient angiogenesis, inflammation, metastasis of a cancerous tissue, excessive or insufficient apoptosis, cardiovascular disease or neurodegenerative diseases. The present sequence is used in the exemplification of the present invention.

Sequence 2097 BP; 483 A; 592 C; 628 G; 394 T; 0 U; 0 Other;

Query Match 99.5%; Score 1035.4; DB 12; Length 2097;  
 Best Local Similarity 99.9%; Pred. No. 3e-294;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGGTAAGG 63  
 Db |||||  
 QY 100 GGTGCTGTGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACCTCCAGCTCGGTAAGG 159  
 Db |||||  
 QY 64 GTGGCTGTGGCTAGGCGCCATTTGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGT 123  
 Db |||||  
 QY 160 GTGGCTGTGGCTAGGCGCCATTTGTGATGGAACAGCGGGAGCAAGTGATCCCCCTGT 219  
 Db |||||  
 QY 124 GTGGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAGACT 183  
 Db |||||  
 QY 220 GTGGCGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAGACT 279  
 Db |||||  
 QY 184 CTCAAATACCACTTTGATGCTTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 243  
 Db |||||  
 QY 280 CTCAAATACCACTTTGATGCTTTCTATGGGAGAGAGTACTCAGCAGACATCTATGCA 339  
 Db |||||  
 QY 244 GTTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGCGCAAGTGCAGTGTGCTGCC 303  
 Db |||||  
 QY 340 GTTCAGTGCAGCCCATCTAAGGCACTTGTGGAAGGCGCAAGTGCAGTGTGCTGCC 399  
 Db |||||  
 QY 304 TATGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCAAGAGCAACTGGG 363  
 Db |||||  
 QY 400 TATGACCCACAGGAGCTGGGAAGACGCAACAATGCTGGGAGCCCAAGAGCAACTGGG 459  
 Db |||||  
 QY 364 GTGATCCCGGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGTCGCGAGGC 423  
 Db |||||  
 QY 460 GTGATCCCGGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGTCGCGAGGC 519  
 Db |||||  
 QY 424 CGGCCATGGGCTTCTGCTCACCATGCTTACCTAGAGATCTACGAGAGAGGTATTA 483  
 Db |||||  
 QY 520 CGGCCATGGGCTTCTGCTCACCATGCTTACCTAGAGATCTACGAGAGAGGTATTA 579  
 Db |||||  
 QY 484 GACCTCTGAGCCCTGCTTGGGAGACTGTTGTAATCCGAGAGAGTGGCGGGGAATATC 543  
 Db |||||  
 QY 580 GACCTCTGAGCCCTGCTTGGGAGACTGTTGTAATCCGAGAGAGTGGCGGGGAATATC 639  
 Db |||||  
 QY 544 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGCGGCACATTC 603  
 Db |||||  
 QY 640 CTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTGTGATTTGAGCGGCACATTC 699  
 Db |||||  
 QY 604 CTGCCAGCCAGTGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC 663  
 Db |||||  
 QY 700 CTGCCAGCCAGTGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCCCGC 759  
 Db |||||  
 QY 664 AGTCATGCTGTGCTCTGCTGAAGTGGACCGAGGAAACGTTTGGCCCCCATTTGCCGAG 723  
 Db |||||  
 QY 760 AGTCATGCTGTGCTCTGCTGAAGTGGACCGAGGAAACGTTTGGCCCCCATTTGCCGAG 819  
 Db |||||  
 QY 724 CGAGAGGGAATCTTACCTGATTTGATTTGGTGGGTGAGAGCAACCGCGGCACAGGC 783  
 Db |||||  
 QY 820 CGAGAGGGAATCTTACCTGATTTGATTTGGTGGGTGAGAGCAACCGCGGCACAGGC 879  
 Db |||||  
 QY 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGGCATCAACACCTCTCTGTTCTCTGGGC 843  
 Db |||||

880 AACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGTTCTGGGC 939  
 Db |||||  
 QY 844 AAAGTGGTAGATGGCTGAATCAGGGCTCCCTCGTGTACTTATCGGGACAGCAGCTC 903  
 Db |||||  
 QY 940 AAAGTGGTAGATGGCTGAATCAGGGCTCCCTCGTGTACTTATCGGGACAGCAGCTC 999  
 Db |||||  
 QY 904 ACTGCGCTATTGAGGAGCTCTCTGGGTGGCTCAGCCACAGTAGTATCTTATTTGCCAAT 963  
 Db |||||  
 QY 1000 ACTGCGCTATTGAGGAGCTCTCTGGGTGGCTCAGCCACAGTAGTATCTTATTTGCCAAT 1059  
 Db |||||  
 QY 964 GCCCTCAGAGAGCTTTCTACCTAGACACAGTTCGCGACTCAACTTTGTCGAGGTCC 1023  
 Db |||||  
 QY 1060 GCCCTCAGAGAGCTTTCTACCTAGACACAGTTCGCGACTCAACTTTGTCGAGGTCC 1119  
 Db |||||  
 QY 1024 AAGGAGGTGATCAATTG 1040  
 Db |||||  
 QY 1120 AAGGAGGTGATCAATCG 1136  
 Db |||||

RESULT 10  
 ADK40995  
 ID ADK40995 standard; DNA; 2099 BP.  
 AC ADK40995;  
 XX AC  
 XX AC  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Novel human kinase gene #15.  
 XX  
 KW cytosolic; immunomodulator; cardiac; neuroprotective; nootropic;  
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
 KW cancer; peripheral nervous system; central nervous system;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; viral infection; prion infection;  
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
 KW attention disorder; cognitive disorder; hypotension; hypertension;  
 KW psychotic disorder; neurological disorder; dyskinesia;  
 KW metabolic disorder; organ transplant rejection; enzyme; gene; ds.  
 XX Homo sapiens.  
 OS  
 XX WO2003057841-A2.  
 PN  
 XX  
 PD 17-JUL-2003.  
 XX  
 XX 31-DEC-2002; 2002WO-US041687.  
 PF  
 XX 31-DEC-2001; 2001US-0343169P.  
 PR  
 XX (GRIG/) GRIGORIEV I V.  
 PA (SUDA/) SUDARSANAM S.  
 XX  
 PI Grigoriev IV, Sudarsanam S;  
 XX  
 XX WPI; 2003-587115/55.  
 DR  
 XX New isolated, enriched or purified nucleic acid molecule encoding a  
 PT kinase polypeptide, useful for treating cancer, immune-related diseases,  
 PT cardiovascular disease, brain or neuronal-associated diseases and  
 PT metabolic disorders.  
 PT  
 XX  
 PS Claim 33; SEQ ID NO 102; 491pp; English.  
 XX  
 CC The invention relates to novel isolated, enriched or purified nucleic acid  
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule  
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and  
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the  
 CC polypeptide in (a), except that it lacks one or more, but not all, of an  
 CC N-terminal domain, C-terminal catalytic domain, a catalytic domain, a C-  
 CC terminal domain, a coiled-coil structure region, a spacer region and a C-

CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules, CC polypeptides, methods and substance are useful for treating cancers, CC immune-related diseases or disorders, cardiovascular disease, brain or CC neuronal-associated diseases, and metabolic disorders. The disorders are CC preferably cancers of the tissues or of hematopoietic origin, diseases of CC the central or peripheral nervous system, Alzheimer's disease, CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, CC viral infections, infections caused by prions, infections caused by CC bacteria, infections caused by fungi, ocular diseases, migraines, pain, CC sexual dysfunction, mood disorders, attention disorders, cognition CC disorders, hypotension, hypertension, psychotic disorders, neurological CC disorders, dyskinesias, metabolic disorders and organ transplant CC rejection. This sequence corresponds to the DNA encoding one of the CC kinase polypeptides of the invention.

XX Sequence 2099 BP; 484 A; 592 C; 628 G; 395 T; 0 U; 0 Other;

Query Match 99.5%; Score 1035.4; DB 10; Length 2099;  
Best Local Similarity 99.9%; Pred. No. 3e-294;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG	63
DB	101	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTAAGG	160
QY	64	GTGGCTGTGCGACTGCGGCCATTTGTGGATGGAACACGCGGAGCAAGTATCCCCCTGT	123
DB	161	GTGGCTGTGCGACTGCGGCCATTTGTGGATGGAACACGCGGAGCAAGTATCCCCCTGT	220
QY	124	GTGCGGGGATGAGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT	183
DB	221	GTGCGGGGATGAGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT	280
QY	184	CTCAATACACAGTTTGATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCA	243
DB	281	CTCAATACACAGTTTGATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCA	340
QY	244	GGTTCACTGACGAGCCATCTCAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTGCTGCC	303
DB	341	GGTTCACTGACGAGCCATCTCAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTGCTGCC	400
QY	304	TATGAGCCACAGAGCTGGGAGAGCGACACAACTGCTGGGAGCCCGAGAGCAACTGGG	363
DB	401	TATGAGCCACAGAGCTGGGAGAGCGACACAACTGCTGGGAGCCCGAGAGCAACTGGG	460
QY	364	GTGATCCCCCGGCTCTCATGACCTCTGTCAGCTCACAAGGAGGAGGAGGTCGCCAGGSC	423
DB	461	GTGATCCCCCGGCTCTCATGACCTCTGTCAGCTCACAAGGAGGAGGAGGTCGCCAGGSC	520
QY	424	CGGCCATGGGCCCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCAGGAGAGGTATTA	483
DB	521	CGGCCATGGGCCCCCTTCTGTCAACCATGCTTACCTAGAGATCTACCAGGAGAGGTATTA	580
QY	484	GACCTCTGGAGCCCTGCTTGGGAGAGCTGGTAATCCGAGAGAGACTGCCGGGGGAATATC	543
DB	581	GACCTCTGGAGCCCTGCTTGGGAGAGCTGGTAATCCGAGAGAGACTGCCGGGGGAATATC	640
QY	544	CTGATTCGCGGCTCTCCAGAGCCCATCAGTACTGCTTGTGATTTTGGAGCGGCACTTC	603
DB	641	CTGATTCGCGGCTCTCTCCAGAGCCCATCAGTACTGCTTGTGATTTTGGAGCGGCACTTC	700
QY	604	CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCCGCTCAACAGCGCTCTCTCCCGC	663
DB	701	CTGCCAGCCAGTCGAAATCGGACTGTAGGAGCCACCCCGCTCAACAGCGCTCTCTCCCGC	760
QY	664	AGTCATGCTGTGCTCTGCTGAAGTGGACACAGCGGGAACGTTTGGCCCCCATTTGCCAG	723
DB	761	AGTCATGCTGTGCTCTGCTGAAGTGGACACAGCGGGAACGTTTGGCCCCCATTTGCCAG	820
QY	724	CCAGAGGGGAAACTCTACTGATTCGATTTGGCTGGGTGAGAGCAACCGGGCGCAGGC	783
DB	821	CCAGAGGGGAAACTCTACTGATTCGATTTGGCTGGGTGAGAGCAACCGGGCGCAGGC	880

QY	784	AACAAGGGGCTTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTTTTGTCTGGGC	843
DB	881	AACAAGGGGCTTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCTCTTTTGTCTGGGC	940
QY	844	AAAGTGGTAGATCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACACAGAGCTC	903
DB	941	AAAGTGGTAGATCGCTGAATCAGGGCCTCCCTCGTGTACCTTATCGGACACAGAGCTC	1000
QY	904	ACTCGCCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTTGGCAACATT	963
DB	1001	ACTCGCCTATTTCAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTTTGGCAACATT	1060
QY	964	GCCCTCTAGAGACGCTTCTACCTAGACACAGTCTCGCAGTCAACTTTGTCGAGGTCC	1023
DB	1061	GCCCTCTAGAGACGCTTCTACCTAGACACAGTCTCGCAGTCAACTTTGTCGAGGTCC	1120
QY	1024	AAGGAGGTGATCAATTG 1040	
DB	1121	AAGGAGGTGATCAATTG 1137	

RESULT 11  
AAFI5853  
ID AAFI5853 standard; cDNA; 2104 BP.  
XX AC AAFI5853;  
XX AC (first entry)  
XX 13-MAR-2001  
XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:288.  
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
KW neuroprotective; cytototoxic; cardioactive; immunomodulatory; muscular;  
KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
KW wound; infectious disease; ss.  
XX Homo sapiens.  
XX WO200055174-A1.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US005988.  
XX 12-MAR-1999; 99US-0124270P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.  
XX Rosen CA, Ruben SM;  
XX WPI; 2000-587513/55.  
XX P-PSDB; AAB56650.  
XX Prostate cancer associated gene sequences, referred to as prostate cancer  
XX antigens, useful for treatment, prevention, and diagnosis of disorders  
XX such as prostate cancer.  
XX Claim 1; Page 805-806; 2338pp; English.  
XX AAFI5566 to AAFI6505 encode the human prostate cancer associated  
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
XX The prostate cancer antigens can have neuroprotective, cytostatic,  
XX cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,  
XX nephrotropic, antiinfective, gynaecological and antibacterial activities,  
XX and can be used in gene therapy. The prostate cancer antigen  
XX polynucleotides may be used for detection of prostate cancer, chromosome  
XX identification, as chromosome markers, and for numerous other diagnostic  
XX or research purposes. The prostate cancer antigens may be used to treat  
XX disorders such as neural, immune, muscular, reproductive,  
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to		
CC AAB57303 represent sequences used in the exemplification of the present		
CC invention		
XX		
SQ	Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;	
Query Match 99.3%; Score 1034; DB 3; Length 2104;		
Best Local Similarity 99.7%; Pred. No. 7.7e-294;		
Matches 1034; Conservative 1; Mismatches 2; Indels 0; Gaps 0;		
QY	4 GGTGGCTGCGCTAAGCAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTAAGG 63	
DB	67 GGTGGCTGTCGGCTAAGCAGATTGGAGCTACTCGTCCACCTCCAGCTCGCGTAAGG 126	
QY	64 GTGGCTGTGCGACTGCGGCCAATTTGTGATGGAAACAGCGGGAGCAAGTGATCCCCCTGT 123	
DB	127 GTGGCTGTGCGACTGCGGCCAATTTGTGATGGAAACAGCGGGAGCAAGTGATCCCCCTGT 186	
QY	124 GTGGGGGCGATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACGAGACT 183	
DB	187 GTGGGGGCGATGGACAGCTGCTCTAGAGATTGCTAACTGGAGAACCCACGAGACT 246	
QY	184 CTCGAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243	
DB	247 CTCGAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 306	
QY	244 GGTTCAGTGCAGCCCATCTTAAGGCATTTGTGAAAGGCGCAGAAATGCCAGTGTCTTGCC 303	
DB	307 GGTTCAGTGCAGCCCATCTTAAGGCATTTGTGAAAGGCGCAGAAATGCCAGTGTCTTGCC 366	
QY	304 TATGACCCACAGGAGCTGGAGAGCAGACACAAATGCTGGCAGCCACAGACCACTTGGG 363	
DB	367 TATGACCCACAGGAGCTGGAGAGCAGACACAAATGCTGGCAGCCACAGACCACTTGGG 426	
QY	364 GTGATCCCGCGGCTCTCATGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGC 423	
DB	427 GTGATCCCGCGGCTCTCATGACCTCTCTGAGCTCACAAGGAGGAGGGTGCAGAGGC 486	
QY	424 CGGCCATGGGCCCTTCTGTCCACATGCTTACCTAGAGATCTACCAGGAGAGGTATTA 483	
DB	487 CGGCCATGGGCCCTTCTGTCCACATGCTTACCTAGAGATCTACCAGGAGAGGTATTA 546	
QY	484 GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCGCGGGGAATATC 543	
DB	547 GACCTCTGGACCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCGCGGGGAATATC 606	
QY	544 CTGATTCCGGGCTCTCCCGAAGCCCATCAGTAGCTTTGTGATTTTGGCGGCACATTC 603	
DB	607 CTGATTCCGGGCTCTCCCGAAGCCCATCAGTAGCTTTGTGATTTTGGCGGCACATTC 666	
QY	604 CTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGCTCAACAGGCTCTCTCCCGC 663	
DB	667 CTGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGCTCAACAGGCTCTCTCCCGC 726	
QY	664 AGTCATGCTGTCTCTGCTCAAGGTGACACAGCGGAACTGTTGGCCCCCAATTTGCCAG 723	
DB	727 AGTCATGCTGTCTCTGCTCAAGGTGACACAGCGGAACTGTTGGCCCCCAATTTGCCAG 786	
QY	724 CGAGGGGAAAACCTTACCTGATTTGACTTTGGCTGGGTGAGAGCAACCGGGGCAAGGC 783	
DB	787 CGAGGGGAAAACCTTACCTGATTTGACTTTGGCTGGGTGAGAGCAACCGGGGCAAGGC 846	
QY	784 AACAGGGCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCTGTTTGTCTCTGGGC 843	
DB	847 AACAGGGCTTCGGCTAAAGAGAGTGGAGCATCAACACCTCTGTTTGTCTCTGGGC 906	
QY	844 AAGTGTGTAGTGCCTGAATCAGGGCCCTCCCTCGTGTATCTTATCGGACAGCAAGCTC 903	
DB	907 AAGTGTGTAGTGCCTGAATCAGGGCCCTCCCTCGTGTATCTTATCGGACAGCAAGCTC 966	
QY	904 ACTCGCTATTGAGAGCTCTGGGTGGCTCAGCCACAGATCTCTTATTCGCAACATT 963	
DB	967 ACTCGCTATTGAGAGCTCTGGGTGGCTCAGCCACAGATCTCTTATTCGCAACATT 1026	

QY	964 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGCAGCTCAACTTGTCTGCCAGGTCC 1023	
DB	1027 GCCCTGAGAGACGCTTCTACCTAGACACAGTCTCGCAGCTCAACTTGTCTGCCAGGTCC 1086	
QY	1024 AAGGAGGTGATCAATTG 1040	
DB	1087 AAGGAGGTGATCAATCG 1103	
RESULT 12		
ACH44493		
ID	ACH44493 standard; cDNA; 464 BP.	
XX	ACH44493;	
AC	ACH44493;	
XX	13-OCT-2003 (first entry)	
DT	XX	
XX	Human foetal brain cDNA #5218.	
DE	XX	
XX	Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;	
KW	genome mapping; biodiversity; genetic disorder.	
KW	XX	
OS	Homo sapiens.	
XX	XX	
XX	US2003073623-A1.	
PN	XX	
XX	17-APR-2003.	
PD	XX	
XX	30-JUL-2001; 2001US-00918995.	
XX	30-JUL-2001; 2001US-00918995.	
PR	XX	
XX	(DRMA/) DRMANAC R T.	
PA	(LABA/) LABAT I.	
PA	(STAC/) STACHE-CRAIN B.	
PA	(DICK/) DICKSON M C.	
PA	(JONE/) JONES L W.	
XX	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;	
PI	XX	
XX	WPI; 2003-615964/58.	
DR	XX	
XX	New polynucleotide sequences obtained from various cDNA libraries, useful	
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene	
PT	mapping, in the recombinant production of protein, or in generating	
PT	antisense DNA or RNA.	
PT	XX	
XX	Claim 1; SEQ ID NO 31705; 44pp; English.	
FS	XX	
XX	The invention relates to an isolated polynucleotide comprising any one of	
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was	
CC	determined by the technique of SBH (sequencing by hybridisation). Also	
CC	included is a purified polypeptide comprising a sequence corresponding to	
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences	
CC	are useful in diagnostics as expressed sequence tags (EST) for	
CC	identifying expressed genes or for physical mapping of the human genome,	
CC	in forensics, in assessing biodiversity, or in identifying mutations	
CC	responsible for genetic disorders and other traits. The nucleotide	
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,	
CC	for chromosome and gene mapping, in the recombinant production of	
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide	
CC	is useful for generating antibodies specific for it. The present sequence	
CC	is one of the 38043 isolated cDNA/EST sequences. Note: the sequence data	
CC	for this patent did not form part of the printed specification, but was	
CC	obtained in electronic format directly from USPTO at	
CC	seqdata.uspto.gov/sequence.html?docID=20030073623	
XX	XX	
SQ	Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;	
Query Match 25.5%; Score 265; DB 9; Length 464;		
Best Local Similarity 77.6%; Pred. No. 1.7e-67;		
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;		



PN US2003194704-A1.  
XX 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
XX (RANK/) RANK D R.  
XX (HANZ/) HANZEL D K.  
XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX Claim 1; SEQ ID NO 24303; 80pp; English.  
XX The invention relates to a nucleic acid probe for measuring human gene  
XX expression, comprising any of the 27,400 fully defined nucleotide  
XX sequences in the specification, or their complements or fragments, and  
XX encoding at least 8 amino acids of any of the 688 amino acid sequences  
XX fully defined in the specification. The probe is a single exon probe that  
XX hybridizes under high stringency conditions to a nucleic acid molecule  
XX expressed in human cells or tissues. Also included are a spatially-  
XX addressed set of single exon nucleic acid probes for measuring human  
XX gene expression (comprising a plurality of single exon nucleic acid  
XX probes cited above, where each of the plurality of probes is separately  
XX and addressably isolatable or amplifiable from the plurality), a single  
XX exon microarray for measuring human gene expression, a method of  
XX measuring human gene expression, a vector comprising the single exon  
XX probe cited above, an ORF-encoded peptide comprising at least 8  
XX contiguous amino acids of any of the above-mentioned amino acid  
XX sequences (optionally with conservative amino acid substitutions), an  
XX isolated antibody that binds specifically to a peptide cited above,  
XX methods of selling and/or licensing single exon probes or microarrays to  
XX a customer desiring to measure gene expression, a method of providing  
XX human gene expression data by subscription, and a computer-readable  
XX storage medium which contains a database having a plurality of records  
XX (each record including data on the expression of a single exon probe  
XX cited above). The probe, methods and apparatus are useful in gene  
XX expression analysis. The probes may be used as tools for surveying  
XX tissues to detect the presence of expressed messages that contain their  
XX specific exon, or in constructing genome-derived single exon microarrays.  
XX In addition, the probes are used in identifying and characterizing gross  
XX alternative splicing events, in detecting and characterizing gross  
XX alterations in the genomic locus that includes their exon, in assessing  
XX smaller genomic alterations, in priming the synthesis of nucleic acids,  
XX or in expressing the ORF-encoded peptide. The present sequence is a human  
XX single exon probe of the invention. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?docID=20030194704  
XX  
XX Sequence 232 BP; 55 A; 64 C; 68 G; 45 T; 0 U; 0 Other;  
Query Match 22.1%; Score 230.4; DB 12; Length 232;  
Best Local Similarity 99.6%; Pred. No. 2e-57; Indels 0; Gaps 0;  
Matches 231; Conservative 0; Mismatches 1;  
QY 688 GTGACACAGCGGAAACGTTTGGCCCATTTTCGCCAGCAGAGGAGGAAACCTCTACCTGATT 747  
DB 1 GTGACACAGCGGAAACGTTTGGCCCATTTTCGCCAGCAGAGGAGGAAACCTCTACCTGATT 60  
QY 748 GACTTGGCTGGGTTCAGAGGACACCGCGCGACAGGCAACAGGGCCCTTCGGCTTAAGAG 807  
DB 61 GACTTGGCTGGGTTCAGAGGACACCGCGCGACAGGCAACAGGGCCCTTCGGCTTAAGAG 120

QY 808 AGTGGAGCATCAACACCTCCCTGTTTCTCTGGGCAAAAGTGGTAGATGCGCTGAATCAG 867  
DB 121 AGTGGAGCATCAACACCTCCCTGTTTCTCTGGGCAAAAGTGGTAGATGCGCTGAATCAG 180  
QY 868 GGCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTGCGCTATTGACGG 919  
DB 181 GGCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCACTGCGCTATTGACGG 232

RESULT 15  
ADM01564  
ID ADM01564 standard; cDNA; 2095 BP.  
XX  
XX ADM01564;  
AC  
XX  
XX 20-MAY-2004 (first entry)  
DT  
XX  
XX Human cDNA of the invention SEQ ID NO:249.  
DB  
XX  
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.  
XX  
XX Homo sapiens.  
OS  
XX  
XX EPI347046-A1.  
PN  
XX  
XX 24-SEP-2003.  
PD  
XX  
XX 12-APR-2002; 2002EP-00008400.  
PF  
XX  
XX 22-MAR-2002; 2002JP-00137785.  
PR  
XX  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
PA  
XX  
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
XX WPI; 2003-723558/69.  
DR  
XX  
XX P-PSDB; ADM04007.  
DR  
XX  
XX New polynucleotides and polypeptides are useful in gene therapy, for  
XX developing a diagnostic marker or medicines for regulating their  
XX expression and activity, or as a target of gene therapy.  
XX  
XX Claim 1; SEQ ID NO 249; 305pp; English.  
XX  
XX The invention relates to a novel human polynucleotide and the encoded  
XX polypeptide. A polynucleotide of the invention may have a use in gene  
XX therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
XX as a primer for synthesizing the polynucleotide or as a probe for  
XX detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
XX useful in gene therapy, for developing a diagnostic marker or medicines  
XX for regulating their expression and activity, or as a target of gene  
XX therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
XX are useful as pharmaceutical agents. The present sequence represents a  
XX cDNA sequence of the invention.  
XX  
XX Sequence 2095 BP; 489 A; 603 C; 658 G; 345 T; 0 U; 0 Other;  
SQ  
Query Match 17.7%; Score 184.6; DB 11; Length 2095;  
Best Local Similarity 53.7%; Pred. No. 1.6e-43;  
Matches 463; Conservative 0; Mismatches 379; Indels 21; Gaps 3;  
QY 190 TACCATGTTGATCGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCGAGTTCA 249  
DB 286 TACCTGTTTCAGCTGGCTTTGACTTCACGCCACCCAGGAGATGTTGATCAGGCCACC 345  
QY 250 GTGACGCCCATCTTAAGGCACCTTGCTGGAAGGCGAAGATGCCAGTGTGCTTGCCTATGCA 309  
DB 346 ACCAAGAGCCTCATCAGGGGGTCTATCTCAGGCTACAAATGCCACTGTCTTTGCTTATGCC 405  
QY 310 CCCACAGAGCTGGGAAGACGACACATGCTGGGCGAGCCAGCAACCTTGGGGTGATC 369

Db 406 CCACAGGCTGTGGGAAACCTACACCATGCTGGGCACAGACAGGAGCCTGGCATCTAT 465  
 QY 370 CCGCGGGCTCTCATGACCTCTCTGACACTCAAAAGGAGGAGGCTGCCAGGGCCGGCCA 429  
 Db 466 GTTCAGACCTCAACGACCTTCCGTGCCATCGAGGAGACCAGCAATGA-----C 516  
 QY 430 TGGGCCCTTCTGTGTCACCATGCTTACCTAGAGATCTACAGGAGAAAGTATTTAGACCTC 489  
 Db 517 ATGGAGTATGAGGTCTCCATGCTCTACCTGGAGATCTACAAATGAGATGATCGGGACCTG 576  
 QY 490 CTGGACCTGCTTCGGGAGCCTGTAATCCGAGAGACTGCGGGGGAAATATCTGTATT 549  
 Db 577 CTGAACCCCTCCCTGGGTACCTGGAGCTCGGGAGGACTTAAGGGGGGTGATCCAGGTG 636  
 QY 550 CCGGGTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTCTTGCCA 609  
 Db 637 GCGGCATCACCGAAGTCTCCACCATCAATGCCAGAGATCATGAGCTGTGATGAAG 696  
 QY 610 GCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCTCTCCGCGAGTCA 669  
 Db 697 GGGAAACCGCAGAGGACCCAGGAGCCACGGCGCCCAACAGAGCTCTCCGCTCCAC 756  
 QY 670 GCTGTCTCTCTGCTAGGTGACAGCGGGAACGTTTGGCCCCATTTGCCA-----G 723  
 Db 757 GGGTACTCAGGTGACCGTGGCCAGCGCGGTCAAGAACATCTTGCAGGAGGTG 816  
 QY 724 CGAGAGGAAACTCTACCTGATTGCTTGGGTGAGAGACAAACCGGCGCACAGGC 783  
 Db 817 CCGCAGGGCGGCTGTTATGATCGACCTGGCTGAGCGCGCTCGGAGACAG 876  
 QY 784 AACAGGGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGC 843  
 Db 877 AATCGTGGCAGCGTATGAAGAGGGGGCCACATCAACCGCTCACTGCTGGCACTGGGC 936  
 QY 844 AAGTGTAGATGCGCTG-----AATCAGGGCTCCCTCGTGTACCTTATCGGACAGC 897  
 Db 937 AATGCTCAACGCCCTGAGCGACAAAGGGTAGCAACAGTACATCAACTATCGGACAGC 996  
 QY 898 AAGCTCACTCGCTATTGAGGACTCTCTGGGTGGCTCAGCCACAGTATCCTTATTGCC 957  
 Db 997 AAGCTCACCGGCTCTGAGGACTCTCTGGAGGAAACAGCCGACAGTATGATCGCT 1056  
 QY 958 AACATTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCC 1017  
 Db 1057 CACATCAGTCTCTCGAGCAGTGCCTTCGAGGAGTCCCGGAACACCTGACCTACGCCGC 1116  
 QY 1018 AGGTCAAGGAGGTGATCAATTG 1040  
 Db 1117 CCGGCCAAGACATTAAGACTAG 1139

Search completed: November 10, 2004, 06:57:56  
 Job time : 541.26 secs

mis Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 13:18:18 ; Search time 555.617 Seconds  
(without alignments)  
10087.639 Million cell updates/sec

Title: US-10-797-893-5  
Perfect score: 1041  
Sequence: 1 atgggtcgctgtcgctgaag.....ccaaggaggatgaattga 1041

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues  
Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW PUB.seq:
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq:
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW PUB.seq:
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW PUB.seq:
13:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
14:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
15:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
16:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
17:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
18:	/cgn2_6/ptodata/1/pubpna/US10_PUB.seq:
19:	/cgn2_6/ptodata/1/pubpna/US11_NEW PUB.seq:
20:	/cgn2_6/ptodata/1/pubpna/US60_NEW PUB.seq:
21:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1041	100.0	1041	US-10-797-893-5	Sequence 5, Appli
2	1038.4	99.8	1464	US-10-797-893-7	Sequence 7, Appli
3	1038	99.7	1115	US-10-797-893-1	Sequence 1, Appli
4	1035.4	99.5	1538	US-10-797-893-3	Sequence 3, Appli
5	1035.4	99.5	2099	US-10-334-143-102	Sequence 102, App
6	1034	99.3	2104	US-09-925-300-288	Sequence 288, App
7	355.4	34.1	24525	US-10-087-192-1180	Sequence 1180, Ap
8	265	25.5	464	US-09-918-995-31705	Sequence 31705, A
9	262.4	25.2	24291	US-10-087-192-1177	Sequence 1177, Ap
10	233.2	22.4	531	US-10-029-386-10603	Sequence 10603, A
11	230.4	22.1	232	US-10-029-386-24303	Sequence 24303, A
12	184.6	17.7	2095	US-10-108-260A-249	Sequence 249, App

13	155.8	15.0	3374	16	US-10-287-226-301	Sequence 301, App
14	136.6	13.1	1152	9	US-09-883-096-6	Sequence 6, Appli
15	136.6	13.1	4108	9	US-09-883-096-1	Sequence 1, Appli
16	135.2	13.0	2675	16	US-10-287-226-305	Sequence 305, App
17	129	12.4	2607	17	US-10-437-963-1460	Sequence 1460, Ap
18	127.6	12.3	482	16	US-10-152-319A-498	Sequence 498, App
19	123.2	11.8	2698	16	US-10-425-114-25129	Sequence 25129, A
20	121.6	11.7	3034	18	US-10-425-115-148256	Sequence 148256, A
21	121.4	11.7	1014	9	US-09-883-096-3	Sequence 3, Appli
22	114.6	11.0	3661	16	US-10-311-642-1	Sequence 1, Appli
23	113.6	10.9	987	16	US-10-311-642-3	Sequence 3, Appli
24	110.8	10.6	4292	16	US-10-287-226-313	Sequence 313, App
25	105.8	10.2	1998	17	US-10-437-963-56548	Sequence 56548, A
26	103.4	9.9	491	9	US-09-960-253-131	Sequence 131, App
27	101	9.7	2264	18	US-10-425-115-1135546	Sequence 135546, A
28	100.2	9.6	3657	16	US-10-336-472-23	Sequence 23, Appli
29	100.2	9.6	3657	16	US-10-236-417-57	Sequence 57, Appli
30	94.2	9.0	4724	15	US-10-172-118-1087	Sequence 1087, Ap
31	94.2	9.0	4724	16	US-10-342-887-1087	Sequence 1087, Ap
32	93.8	9.0	3258	16	US-10-302-173-586	Sequence 586, App
33	93.4	9.0	2217	15	US-10-104-047-201	Sequence 201, App
34	90.4	8.7	4913	15	US-10-133-937-34	Sequence 34, Appli
35	90.4	8.7	4913	16	US-10-159-563-34	Sequence 34, Appli
36	88.8	8.5	1295	15	US-10-369-493-27001	Sequence 27001, A
37	88.4	8.5	3159	15	US-10-104-047-762	Sequence 762, App
38	88	8.5	3390	16	US-10-425-114-22135	Sequence 22135, A
39	88	8.5	3407	18	US-10-739-930-2359	Sequence 2359, A
40	88	8.5	3433	18	US-10-425-115-46568	Sequence 46568, A
41	86.2	8.3	6927	15	US-10-172-118-1053	Sequence 1053, Ap
42	86.2	8.3	6927	16	US-10-342-887-1053	Sequence 1053, Ap
43	81.8	7.9	379	16	US-10-424-599-111387	Sequence 111387, A
44	80.8	7.8	4037	16	US-10-287-226-327	Sequence 327, App
45	79.2	7.6	1032	14	US-10-159-151-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1  
US-10-797-893-5  
; Sequence 5, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Berand, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Human  
US-10-797-893-5

Query Match	100.0%	Score 1041;	DB 17;	Length 1041;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1041;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGTGCTGTCGGCTAAGCAAGATTGGAGTCTCGTCTCGTCTCGTCTCGGTA	60	
Db	1	ATGGGTGCTGTCGGCTAAGCAAGATTGGAGTCTCGTCTCGTCTCGGTA	60	
QY	61	AGGGTGGTGTGCGACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTATCCCCC	120	
Db	61	AGGGTGGTGTGCGACTCGCGCCATTGTGGATGGAACAGCGGAGCAAGTATCCCCC	120	

```

QY 121 TGTGTGCGGGGATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAG 180
DB 121 TGTGTGCGGGGATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAG 180
QY 181 ACTCTCAATACACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
DB 181 ACTCTCAATACACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
QY 241 GCAGTTTCAGTACGCCCATCTAAGGCATTTGCTTGGAAAGGCAGAAATGCCAGTGTGTT 300
DB 241 GCAGTTTCAGTACGCCCATCTAAGGCATTTGCTTGGAAAGGCAGAAATGCCAGTGTGTT 300
QY 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAAATGCTGGGAGAGAGGAGGAGGAG 360
DB 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAAATGCTGGGAGAGAGGAGGAGGAG 360
QY 361 GGGGTGATGTCGGGGCTCTCATGACCTCTCTGAGCTCACAAAGGAGAGAGGAGGAGG 420
DB 361 GGGGTGATGTCGGGGCTCTCATGACCTCTCTGAGCTCACAAAGGAGAGAGGAGGAGG 420
QY 421 GCGCGGCATGAGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
DB 421 GCGCGGCATGAGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
QY 481 TTAGACCTCTGACCCCTCTTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGGCGAC 540
DB 481 TTAGACCTCTGACCCCTCTTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGGCGAC 540
QY 541 ATCTCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGGCGAC 600
DB 541 ATCTCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGGCGAC 600
QY 601 TTCCTGCGAGCCAGTGCAGAACTCGAATCTGTAGAGAGCCCGGCTCAACAGGCTCTCTCC 660
DB 601 TTCCTGCGAGCCAGTGCAGAACTCGAATCTGTAGAGAGCCCGGCTCAACAGGCTCTCTCC 660
QY 661 GCGAGTCATGCTGCTCTGCTCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 661 GCGAGTCATGCTGCTCTGCTCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 721 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 721 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 GGCACACAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 781 GGCACACAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
QY 841 GGCACAGTGTAGATGCGCTGAATCAGGGCTCCCTCGGTGACCTTATCGGGAGAGCAAG 900
DB 841 GGCACAGTGTAGATGCGCTGAATCAGGGCTCCCTCGGTGACCTTATCGGGAGAGCAAG 900
QY 901 CTCACCTCGCTATTCAGGAGCTCTCTGGGTGGCTCAGGCCACAGATCTCTTATTCGCAAC 960
DB 901 CTCACCTCGCTATTCAGGAGCTCTCTGGGTGGCTCAGGCCACAGATCTCTTATTCGCAAC 960
QY 961 ATTGCCCTGAGAGAGGCTTCTACCTAGACAGAGTCTCGGACTCAACTTTCGCGAGG 1020
DB 961 ATTGCCCTGAGAGAGGCTTCTACCTAGACAGAGTCTCGGACTCAACTTTCGCGAGG 1020
QY 1021 TCCAGGAGGATCAATTTGA 1041
DB 1021 TCCAGGAGGATCAATTTGA 1041

```

RESULT 2

```

US-10-797-893-7
; Sequence 7, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use

```

```

; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/10797,893
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
; US-10-797-893-7

```

```

Query Match 99.8%; Score 1038.4; DB 17; Length 1464;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1039; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60
DB 1 ATGGGTGCTGCTCGGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCAGCTCGCGTA 60
QY 61 AGGGTGGCTGTGCCACTGCGGCCCATTTGTGATGGAAACAGCGGGAGCAAGTGATCCCCC 120
DB 61 AGGGTGGCTGTGCCACTGCGGCCCATTTGTGATGGAAACAGCGGGAGCAAGTGATCCCCC 120
QY 121 TGTGTGCGGGGATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAG 180
DB 121 TGTGTGCGGGGATGAGCAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCCAGGAG 180
QY 181 ACTCTCAATACACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
DB 181 ACTCTCAATACACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGACATCTAT 240
QY 241 GCAGTTTCAGTACGCCCATCTAAGGCATTTGCTTGGAAAGGCAGAAATGCCAGTGTGTT 300
DB 241 GCAGTTTCAGTACGCCCATCTAAGGCATTTGCTTGGAAAGGCAGAAATGCCAGTGTGTT 300
QY 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAAATGCTGGGAGAGAGGAGGAGGAG 360
DB 301 GCCTATGACCCACAGGAGCTGGGAGAGCAGCACAAATGCTGGGAGAGAGGAGGAGGAG 360
QY 361 GGGGTGATGTCGGGGCTCTCATGACCTCTGAGCTCACAAAGGAGAGAGGAGGAGGAGG 420
DB 361 GGGGTGATGTCGGGGCTCTCATGACCTCTGAGCTCACAAAGGAGAGAGGAGGAGGAGG 420
QY 421 GCGCGGCATGAGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
DB 421 GCGCGGCATGAGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTA 480
QY 481 TTAGACCTCTGAGACCTCTCGGGAGACCTGGTAACTCGGAGAGACTCGCGGGGAGAT 540
DB 481 TTAGACCTCTGAGACCTCTCGGGAGACCTGGTAACTCGGAGAGACTCGCGGGGAGAT 540
QY 541 ATCTCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGGCGAC 600
DB 541 ATCTCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGGCGAC 600
QY 601 TTCCTGCGAGCCAGTGCAGAACTCGAATCTGTAGAGAGCCCGGCTCAACAGGCTCTCTCC 660
DB 601 TTCCTGCGAGCCAGTGCAGAACTCGAATCTGTAGAGAGCCCGGCTCAACAGGCTCTCTCC 660
QY 661 GCGAGTCATGCTGCTCTGCTCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 661 GCGAGTCATGCTGCTCTGCTCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 721 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 721 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 GGCACACAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
DB 781 GGCACACAGGGCTTCGGCTAAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840

```

Db 781 GGCACAAAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGTCCTG 840  
Qy 841 GGCACAAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCCGGACAGCAAG 900  
Db 841 GGCACAAAGTGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCCGGACAGCAAG 900  
Qy 901 CTCACTCGCCCTATTGCAAGACTCTCTGGGTGGCTCAGCCCAACAGATGCTTATGCAAC 960  
Db 901 CTCACTCGCCCTATTGCAAGACTCTCTGGGTGGCTCAGCCCAACAGATGCTTATGCAAC 960  
Qy 961 ATTGCCCTCAGAGAGCGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020  
Db 961 ATTGCCCTCAGAGAGCGCTTCTACCTAGACACAGTCTCCGACCTCAACTTTGCTGCCAGG 1020  
Qy 1021 TCCAAGGAGGTGATCAATTG 1040  
Db 1021 TCCAAGGAGGTGATCAATCG 1040  
RESULT 3  
US-10-797-893-1  
; Sequence 1, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1115  
; TYPE: DNA  
; ORGANISM: Human  
US-10-797-893-1  
Query Match 99.7%; Score 1038; DB 17; Length 1115;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1038; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 GTTCGCTGTCGGCTAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGTCGCGTAAGG 63  
Db 78 GTTCGCTGTCGGCTAGCAAGATTGGAGCTACTCTGCTCCACCTCCAGTCGCGTAAGG 137  
Qy 64 GTTCGCTGTCGACTGCGGCCATTGTGGATGGAAACAGCGGAGCAAGTATCCCTGT 123  
Db 138 GTTCGCTGTCGACTGCGGCCATTGTGGATGGAAACAGCGGAGCAAGTATCCCTGT 197  
Qy 124 GTTCGCGGCGATGACAGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 183  
Db 198 GTTCGCGGCGATGACAGCTCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGACT 257  
Qy 184 CTCAATACCAAGTTGATGCTCTATGCGGAGAGAGTACTCAGCAGGACATCTATGCA 243  
Db 258 CTCAATACCAAGTTGATGCTCTATGCGGAGAGAGTACTCAGCAGGACATCTATGCA 317  
Qy 244 GGTTCAGTCAGCCCACTTAAGCACTTCTCGAAGGGCAGAAATCCAGTGTGCTTGC 303  
Db 318 GGTTCAGTCAGCCCACTTAAGCACTTCTCGAAGGGCAGAAATCCAGTGTGCTTGC 377  
Qy 304 TATGGACCCACAGGAGTGGGAAGAGCGCACAAATGCTGGGACAGCCACAGCAACTGGG 363  
Db 378 TATGGACCCACAGGAGTGGGAAGAGCGCACAAATGCTGGGACAGCCACAGCAACTGGG 437  
Qy 364 GTGATCCCGGGGCTCTCATGACCTCTGACCTCAGTCAAGGAGGAGGTGCCGAGGC 423  
Db 438 GTGATCCCGGGGCTCTCATGACCTCTGACCTCAGTCAAGGAGGAGGTGCCGAGGC 497

Qy 424 CGSCCATGGGCCCTTTCTGTACCATGTCTTACCTAGATCTTACAGGAGAAAGGTATTA 483  
Db 498 CGSCCATGGGCCCTTTCTGTACCATGTCTTACCTAGATCTTACAGGAGAAAGGTATTA 557  
Qy 484 GACCTCTGGAACCTCTGCTCGGAGAGACCTGGTAATCCGAGAAGACTGCGCGGGGAATATC 543  
Db 558 GACCTCTGGAACCTCTGCTCGGAGAGACCTGGTAATCCGAGAAGACTGCGCGGGGAATATC 617  
Qy 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC 603  
Db 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGAGCGCACTTC 677  
Qy 604 CTGCCACGACAGTCCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCCGC 663  
Db 678 CTGCCACGACAGTCCGAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCCGC 737  
Qy 664 AGTCATGCTGTCTCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCCCCCATTTGCGCAG 723  
Db 738 AGTCATGCTGTCTCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCCCCCATTTGCGCAG 797  
Qy 724 CGAGAGGGGAAACTCTACCTGATTGACTTGGCTGGGTTCAGAGGACAAACCGCGCACAGGC 783  
Db 798 CGAGAGGGGAAACTCTACCTGATTGACTTGGCTGGGTTCAGAGGACAAACCGCGCACAGGC 857  
Qy 784 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGCG 843  
Db 858 AACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTCGGCG 917  
Qy 844 AAGTGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGAGCAGCAAGCTC 903  
Db 918 AAGTGTAGATGCGCTGAATCAGGGCCCTCCCTGCTGTACCTTATCGGAGCAGCAAGCTC 977  
Qy 904 ACTCGCTATTGCAAGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGCGCAACTT 963  
Db 978 ACTCGCTATTGCAAGACTCTCTGGGTGGCTCAGCCACAGTATCTTATGCGCAACTT 1037  
Qy 964 GCGCTTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTGTGCGCAGGTCC 1023  
Db 1038 GCGCTTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTGTGCGCAGGTCC 1097  
Qy 1024 AAGAGGTGATCAATTGA 1041  
Db 1098 AAGAGGTGATCAATTGA 1115

## RESULT 4

US-10-797-893-3  
; Sequence 3, Application US/10797893  
; Publication No. US20040142397A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/797,893  
; CURRENT FILING DATE: 2004-03-09  
; PRIOR APPLICATION NUMBER: US/09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/597,292  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1538  
; TYPE: DNA  
; ORGANISM: Human  
US-10-797-893-3

Query Match 99.5%; Score 1035.4; DB 17; Length 1538;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 63  
 Db |||||  
 QY 78 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 137  
 Db |||||  
 QY 64 GTGGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 123  
 Db |||||  
 QY 138 GTGGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 197  
 Db |||||  
 QY 124 GTGGGGGATGAGCAGCTGTCTCTAGAGATTGTAACTGAGAGAACCCAGAGAGACT 183  
 Db |||||  
 QY 198 GTGGGGGATGAGCAGCTGTCTCTAGAGATTGTAACTGAGAGAACCCAGAGAGACT 257  
 Db |||||  
 QY 184 CTCAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCAGAGAGACTATGCA 243  
 Db |||||  
 QY 258 CTCAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCAGAGAGACTATGCA 317  
 Db |||||  
 QY 244 GGTTCAGTGCAGCCCATCTAAGGCATCTTGTGGAAGGCGAAGTCCAGTGTGCTGCC 303  
 Db |||||  
 QY 318 GGTTCAGTGCAGCCCATCTAAGGCATCTTGTGGAAGGCGAAGTCCAGTGTGCTGCC 377  
 Db |||||  
 QY 304 TATGAGCCACAGGAGCTGGAGAGCGACACAAATGCTGGGAGAGAGAGCAACCTGGG 363  
 Db |||||  
 QY 378 TATGAGCCACAGGAGCTGGAGAGCGACACAAATGCTGGGAGAGAGAGCAACCTGGG 437  
 Db |||||  
 QY 364 GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAGGAGAGAGGTCGCGAGGGC 423  
 Db |||||  
 QY 438 GTGATCCCGGGCTCTCATGACCTCTGACAGCTCACAAGGAGAGAGGTCGCGAGGGC 497  
 Db |||||  
 QY 424 CGGCCATGCGGCTCTTCTGTCAACCATCTTACCTAGAGATCTACAGAGAGAGTATTA 483  
 Db |||||  
 QY 498 CGGCCATGCGGCTCTTCTGTCAACCATCTTACCTAGAGATCTACAGAGAGAGTATTA 557  
 Db |||||  
 QY 484 GACCTCTGGAACCTGCTTGGGAGAGCTGTGTAATCCGAGAGAGCTGCGGGGGAATATC 543  
 Db |||||  
 QY 558 GACCTCTGGAACCTGCTTGGGAGAGCTGTGTAATCCGAGAGAGCTGCGGGGGAATATC 617  
 Db |||||  
 QY 544 CTGATTCGGGCTCTCCAGAGCCCATCAGTGTGCTGATTTGAGCGGAGCTTC 603  
 Db |||||  
 QY 618 CTGATTCGGGCTCTCCAGAGCCCATCAGTGTGCTGATTTGAGCGGAGCTTC 677  
 Db |||||  
 QY 604 CTGCGCAGCAGTCTGAATCGGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 663  
 Db |||||  
 QY 678 CTGCGCAGCAGTCTGAATCGGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC 737  
 Db |||||  
 QY 664 AGTCATGCTGTCTGTCTGAGTGAACAGCGGGAACGTTTGGCCCAATTCGCCAG 723  
 Db |||||  
 QY 738 AGTCATGCTGTCTGTCTGAGTGAACAGCGGGAACGTTTGGCCCAATTCGCCAG 797  
 Db |||||  
 QY 724 CGAGAGGGAACCTCTACCTGATTTGACTTGGCTGAGAGCAACCGGCGCAGAGC 783  
 Db |||||  
 QY 798 CGAGAGGGAACCTCTACCTGATTTGACTTGGCTGAGAGCAACCGGCGCAGAGC 857  
 Db |||||  
 QY 784 AACAGGGGCTTTCGGCTAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTTGGGC 843  
 Db |||||  
 QY 858 AACAGGGGCTTTCGGCTAAGAGAGTGGAGCCATCAACACCTCTGTTTGTCTTGGGC 917  
 Db |||||  
 QY 844 AAGTGTGATGAGTGTGATGAGGCTCTGCTGTGATGATGATGATGATGATGATGATGAT 903  
 Db |||||  
 QY 918 AAGTGTGATGAGTGTGATGAGGCTCTGCTGTGATGATGATGATGATGATGATGATGAT 977  
 Db |||||  
 QY 904 ACTCGCTTATGAGGAGCTCTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 963  
 Db |||||  
 QY 978 ACTCGCTTATGAGGAGCTCTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 1037  
 Db |||||  
 QY 964 GCGCCCTGAGAGAGCTTCTACCTAGACAGCTCTCGGACTCAACTTGTCTCGGAGTCC 1023  
 Db |||||  
 QY 1038 GCGCCCTGAGAGAGCTTCTACCTAGACAGCTCTCGGACTCAACTTGTCTCGGAGTCC 1097  
 Db |||||  
 QY 1024 AAGGAGGTGATCAATTG 1040  
 Db |||||  
 QY 1098 AAGGAGGTGATCAATCG 1114  
 Db |||||

RESULT 5  
 US-10-334-143-102  
 ; Sequence 102, Application US/10334143  
 ; Publication No. US20040009549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
 ; APPLICANT: SUDARSANAM, SUCHA  
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
 ; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
 ; FILE REFERENCE: 038602/1543  
 ; CURRENT APPLICATION NUMBER: US/10/334,143  
 ; CURRENT FILING DATE: 2002-12-31  
 ; PRIOR APPLICATION NUMBER: 60/343,169  
 ; PRIOR FILING DATE: 2001-12-31  
 ; NUMBER OF SEQ ID NOS: 207  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 102  
 ; LENGTH: 2099  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-334-143-102  
 Query Match 99.5%; Score 1035.4; DB 16; Length 2099;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 63  
 Db |||||  
 QY 101 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 160  
 Db |||||  
 QY 64 GTGGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 123  
 Db |||||  
 QY 161 GTGGCTGTGCGCTAAGCAAGATTGGAGTACTCGTGTCCACCTCCAGCTCGCGTAAGG 220  
 Db |||||  
 QY 124 GTGGGGGATGAGCAGCTGTCTCTAGAGATTGTAACTGAGAGAACCCAGAGAGACT 183  
 Db |||||  
 QY 221 GTGGGGGATGAGCAGCTGTCTCTAGAGATTGTAACTGAGAGAACCCAGAGAGACT 280  
 Db |||||  
 QY 184 CTCAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCAGAGAGAGTATGCA 243  
 Db |||||  
 QY 281 CTCAATACCAAGTTGATGCTCTATGGGAGAGAGTACTCAGAGAGAGTATGCA 340  
 Db |||||  
 QY 244 GTTCAGTGCAGCCCATCTAAGGCATCTTGTGGAAGGCGAAGTCCAGTGTGCTGCC 303  
 Db |||||  
 QY 341 GTTCAGTGCAGCCCATCTAAGGCATCTTGTGGAAGGCGAAGTCCAGTGTGCTGCC 400  
 Db |||||  
 QY 304 TATGAGCCACAGGAGCTGGAGAGCGACACAAATGCTGGGAGAGAGCAACCTGGG 363  
 Db |||||  
 QY 401 TATGAGCCACAGGAGCTGGAGAGCGACACAAATGCTGGGAGAGAGCAACCTGGG 460  
 Db |||||  
 QY 364 GTGATCCCGGGCTCTCATGGAACCTCTGAGAGCTCAAGGAGAGAGGTCGCGAGGGC 423  
 Db |||||  
 QY 461 GTGATCCCGGGCTCTCATGGAACCTCTGAGAGCTCAAGGAGAGAGGTCGCGAGGGC 520  
 Db |||||  
 QY 424 CGGCCATGCGGCTCTTCTGTCAACCATCTTACCTAGAGATCTACAGAGAGAGTATTA 483  
 Db |||||  
 QY 521 CGGCCATGCGGCTCTTCTGTCAACCATCTTACCTAGAGATCTACAGAGAGAGTATTA 580  
 Db |||||  
 QY 484 GACCTCTGGAACCTGCTTGGGAGAGCTGTGTAATCCGAGAGAGCTGCGGGGGAATATC 543  
 Db |||||  
 QY 581 GACCTCTGGAACCTGCTTGGGAGAGCTGTGTAATCCGAGAGAGCTGCGGGGGAATATC 640  
 Db |||||  
 QY 544 CTGATTCGGGCTCTTCTGAGAGAGCCATCAGTGTGCTGATTTTGTGAGCGGAGCTTC 603  
 Db |||||  
 QY 641 CTGATTCGGGCTCTTCTGAGAGAGCCATCAGTGTGCTGATTTTGTGAGCGGAGCTTC 700  
 Db |||||  
 QY 604 CTGCGCAGCAGTCTGAATCGGAGCTGTAGAGCCACCGGCTCAACAGCGCTCTCCCGC 663  
 Db |||||  
 QY 701 CTGCGCAGCAGTCTGAATCGGAGCTGTAGAGCCACCGGCTCAACAGCGCTCTCCCGC 760  
 Db |||||  
 QY 664 AGTCATGCTGTCTCTGCTCAAGGTGAGCAGCGGGGAACGTTTGGCCCAATTCGCCAG 723  
 Db |||||  
 QY 761 AGTCATGCTGTCTCTGCTCAAGGTGAGCAGCGGGGAACGTTTGGCCCAATTCGCCAG 820  
 Db |||||



```

; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1180
; LENGTH: 24525
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1180

```

```

Query Match      34.1%; Score 355.4; DB 13; Length 24525;
Best Local Similarity 84.8%; Pred. No. 2.3e-103;
Matches 445; Conservative 0; Mismatches 1; Indels 79; Gaps 1;

QY 476 AGGTATTAGACCTCTCGGAGACCTGCTCGGAGACCTGTAATCCGAGAAGACTGCCGGG 535
DB 2313 AGGTATTAGACCTCTCGGAGACCTGCTCGGAGACCTGTAATCCGAGAAGACTGCCGGG 2372
QY 536 GGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGC 595
DB 2373 GGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGAGC 2432
QY 596 GGCACCTCTGCGACGAGTGAATCGGACTGTAGGAGCCACCGGCTCAACCGCGCT 655
DB 2433 GGCACCTCTGCGACGAGTGAATCGGACTGTAGGAGCCACCGGCTCAACCGCGCT 2492
QY 656 CTTCCCGCAGTCATCTGCTGCTCTGCTCA-----AGGTGACCCAG 696
DB 2493 CTTCCCGCAGTCATCTGCTGCTCTGCTCAAGGTGAGCGCCGAGAGGAGACCT 2552
QY 696 -----AGGTGACCCAG 696
DB 2553 GGAAGCCAGGAGCTGAGCTAAGCAGGAGACCTTTGTTCTTACCCCGAGTGGACCCAG 2612
QY 697 CGGGAACCTTTGGCCCAATTTGCGCAGCGAGAGGGAACCTTACCTGATTTGCTTGGCT 756
DB 2613 CGGGAACCTTTGGCCCAATTTGCGCAGCGAGAGGGAACCTTACCTGATTTGCTTGGCT 2672
QY 757 GGGTCAGAGGACAAACCGCGCACAGGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCC 816
DB 2673 GGGTCAGAGGACAAACCGCGCACAGGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCC 2732
QY 817 ATCAACACCTCCCTGTTTGTCTCGGCAAGTGTAGTGCCTGAATTCAGGGCTCCCT 876
DB 2733 ATCAACACCTCCCTGTTTGTCTCGGCAAGTGTAGTGCCTGAATTCAGGGCTCCCT 2792
QY 877 CGGTACCTTATCGGAGCAGCAAGCTCACTCGCTATTGCAAGGAC 921
DB 2793 CGGTACCTTATCGGAGCAGCAAGCTCACTCGCTATTGCAAGGTC 2837

```

```

RESULT 8
US-09-918-995-31705
; Sequence 31705, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31705
; LENGTH: 464
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(464)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-31705

Query Match      25.5%; Score 265; DB 10; Length 464;
Best Local Similarity 77.6%; Pred. No. 1e-74;
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 126 GCGGGGATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCT 185
DB 37 GCGTGGAAATCAATATCGGGTCGATAGTGTAAACGCGCCCAACCAACCCGATAATCT 96
QY 186 CAAATACCAAGTTTATGCTTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGG 245
DB 97 CCGTTACTGTGCTTGCCTTTGACAGCCATAGGAGACTGCGCTGACATGTACCTTCC 156
QY 246 TTCAGTGCAGCCCATCTTAAGGCACTTCTGGAAGGGGAGAGTATGCGAGTGTGCTTGCCTA 305
DB 157 TTGACCGTGGACCGTGCCTTTCGTAATAAAGCGAGGAGCTGGCCTA 216
QY 306 TGGACCCACAGAGCTGCGAAGACGCAACAATCTGGGAGCCGAGAGCAACCTGGGGT 365
DB 217 CGGTGCTTGGAGAGCTGGCATTAC-ATCTGATGCTTTGGCAGCCGAGAGCAACCTGGGT 275
QY 366 GATCCCGGGCTCTCATGAGACCTCTGCTGAGCTCACAAGGAGGAGGAGTGCAGGCGC 425
DB 276 GATCCCGGGCTCTCATGAGACCTCTGCTGAGCTCACAAGGAGGAGGAGTGCAGGCGC 335
QY 426 GCGATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCAAGAGAGGTATTAGA 485
DB 336 GCGATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCAAGAGAGGTATTAGA 395
QY 486 CCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAAGACTGCGGGGGAATATCCT 545
DB 396 CCTCTGAGCCCTGCTTGGGAGACCTTGGNAATCCGAGAAGACTGCGGGGGAATATCCT 455
QY 546 GATTCCGGG 554
DB 456 GATTCCGGG 464

```

```

RESULT 9
US-10-087-192-1177
; Sequence 1177, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1177
; LENGTH: 24291
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(24291)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1177

```

```

Query Match      25.2%; Score 262.4; DB 13; Length 24291;

```



```

; TITLE OF INVENTION: No. US2004005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2095
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-108-260A-249

Query Match          17.7%  Score 184.6;  DB 16;  Length 2095;
Best Local Similarity 53.7%;  Pred. No. 1.3e-48;
Matches 463;  Conservative 0;  Mismatches 379;  Indels 21;  Gaps 3;

QY 190 TACAGTTTATGCTCTTATGCGGAGAGGAGTACTCAGCAGACATCTATGTCAGGTTC 249
DB 286 TACCTGTTCGACGTGGCTTTGACTTCACCGCCACCCAGGAGATGTTGATCAGGCCACC 345
QY 250 GTCCAGCCCATCTAAGGCATCTGCTGGAGGCGAAGATGCGAGTGTGCTGCTATGGA 309
DB 346 ACCAAGAGCCTCATCAGAGGCGTCTATCAGGCTACAATGCCACTGTCTTTCCTATGCG 405
QY 310 CCCACAGGAGCTGGGAGAGCGACACAATGCTGGGAGCCCGAGAGCAACCTGGGTGATC 369
DB 406 CCCACAGGCTGGGAGAGCGACACAATGCTGGGAGCCCGAGAGCAACCTGGGTGATC 465
QY 370 CCGCGGGCTCTCATGACCTCTCTGACCTCACAAAGGAGGAGGGTCCGAGGGCCCGGCCA 429
DB 466 GTTCAGACCTCAACGACCTTTCCTGCTCATCGAGGAGCAGCAATGA-----C 516
QY 430 TGGGGCTTCTGTACCATGCTTACCTAGAGATCTACAGGAGAGGATTTAGACCTC 489
DB 517 ATGAGTATGAGTGTCTTCTGCTCTTCTGAGATCTACAATGAGATGATCGGGACCTG 576
QY 490 CTGACCTCTCTGCGGAGACCTGTGTAATCCGAGAGAGCTGCGGGGGAATATCTTGATT 549
DB 577 CTGAACCTCTCTGCGGAGACCTGTGTAATCCGAGAGAGCTGCGGGGGAATATCTTGATT 636
QY 550 CCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTCTTCGCCA 609
DB 637 GCGGCGATCACCAGAGCTCTCCACCATGCAAGGAGATCATGCGCTGCTGATGAAG 696
QY 610 GCGAGTCAATCGGAGCTAGAGGACCCCGCTCAACGAGGCTCTCTCCGAGTAT 669
DB 697 GGGAACTGCGAGAGGACCCAGAGGACCCCGCGCCCAACGAGAGCTCTCTCCGCTCCCA 756
QY 670 GCTGTCTCTGCTCAAGGTGACGAGCGGAACTGTTGGCCCCCATTTCCGCA-----G 723
DB 757 GCGGTACTGAGGTGACCTGCGGAGCGGAGCGGAGCTCAAGACATCTTGAGGAGGTG 816
QY 724 CGAGAGGGAATCTTACTGATGATGATGCTGCTGAGGAGCAACCGGCGCAGAGC 783
DB 817 CGGAGGCGGCTGCTGATGATGATGCTGCTGAGGAGCAACCGGCGCAGAGCAG 876
QY 784 AACAGGGCTTGGGTAAGAGAGTGGAGCCATCAACACCTCTCTGCTGCTGCGC 843
DB 877 AATCGTGGGAGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 936
QY 844 AAGTGGTATGATGCTG-----AATCAGGGGCTCTCTGCTGATCTTATCGGAGCAGC 897
DB 937 AACTGATCAACGCTCTGAGCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 996
QY 898 AGCTCACTGCGCTATGTCAGGAGCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
DB 997 AAGCTCAGCGGCTCTGAGGAGCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1056
QY 958 AACATTGCCCTGAGAGGCTTCTTACCTAGACAGCTCTCGGAGTCACTTCTGCTGCC 1017
DB 1057 CACATGAGTCTGCGAGGAGTGTCTTTCGAGGAGTGTCTGAGGAGTGTCTGAGGAGTGTCT 1116
QY 1018 AGGTCCAGGAGGTGATCAATTG 1040

```

```

DB 1117 CCGGCCAGACATTAAGACTAG 1139

RESULT 13
US-10-287-226-301
; Sequence 301, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zethusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraSeqList version 0.1

```



SEQ ID NO 301  
LENGTH: 3374  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (7)..(3351)  
US-10-287-226-301

Query Match 15.0%; Score 155.8; DB 16; Length 3374;  
Best Local Similarity 54.0%; Pred. No. 2.9e-39;  
Matches 394; Conservative 0; Mismatches 317; Indels 18; Gaps 3;  
QY 321 TGGGAAGAGCCACACAAATGCTGGGCGAGCCGAGAGCAACCTGGGGTATCCCGGGGCTCT 380  
DB 612 TGGGAAGAGCCACACAAATGCTGGGCGAGCCGAGAGCAACCTGGGGTATCCCGGGGCTCT 671  
QY 381 CATGGACCTCTCGAGCTCAAGAGGAGGAGGTGGGAGGCGGCGGCAATGGGCCCTTTC 440  
DB 672 CAACGACCTCTTCGTCGCCATCGAGGAGACCAAGCAATGA-----CATGGAGTATGA 722  
QY 441 TGTCCACATGCTTACTAGAGATCTACCAAGGAGAGGTATTAGACCTCTGACCCCTGC 500  
DB 723 GGTCTCCATGCTCTACTGGAGATCTACATGAGATGATCCGGGACCTGCTGAACCCCTC 782  
QY 501 TTGGGAGACCTGGTAAATCCGAGAAGACTGCCGGGGAAATATCTTGATTCGGGGTCTCTC 560  
DB 783 CCTGGGCTACCTGGAGCTGCGGAGGACTCTAAGGGGGTATCCAGGTGGCGGCATCAC 842  
QY 561 CCAAGAGCCCATCAGTAGCTTTCTGATTTGAGCGGCACTTCTGCGAGCCAGTGCAGAA 620  
DB 843 CGAAGTCTCCACATCAATGTCGAAGGAGATCATGACGTGCTGATGAAGGGGAACCGGCA 902  
QY 621 TCGGAGCTGAGGAGCCACCGGCTCAACAGCGCTCTCCCGAGCTCATGCTGTGCTCT 680  
DB 903 GAGGACCCAGGAGCCACCGGCCCCCAACAGAGCTCTCCCGCTCCCGAGCGGTACTGCA 962  
QY 681 GGTCAAGGTGGACAGCGGGAAAGCTTT---GGCCCCATTTCCGCGAGGAGGGGAAACT 737  
DB 963 GGTGACCGTGGCCAGCGAGCGGGTCAAGAACATCTTGCAGGAGCGCGGCGGCT 1022  
QY 738 CTACCTGATTGACTCTGGTGGGTGAGAGCAACCGCGGACAGGCAACAGGGCGCTTCG 797  
DB 1023 GTTCATGATCGACTGGCTGGCTCAGAGCGGCTCGCAGACACAGAAATCGTGGGCGAG 1082  
QY 798 GCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTCTGCGGCAAGTGTAGATGC 857  
DB 1083 TATGAAGGAGGGGGCCCATCAACCGCTCACTGCTGGCACTGGGCAATGCAATCAACGC 1142  
QY 858 GCTGAATCAGG-----GCTCCTCGTGTACCTTATCGGAGCAGCAAGTCACTCGCCT 911  
DB 1143 CTGAGCGACAGGGGTAGCAACAAGTACATCAACTATCGGACAGCAAGTCAACCGGT 1202  
QY 912 ATTCAGGACTCTCTGGTGGGTGAGCCACAGTATCTTATTCGCAACATTGCCCTGA 971  
DB 1203 CTTGAAGGACTCTCTGGGAGGAAACAGCGGACAGTATGCTCATCATAGTCTCTG 1262  
QY 972 GAGACCTTCTACCTAGACAGTCTCCGACTCAACTTTGCTCCAGGTCCAGAGGT 1031  
DB 1263 GAGCAGTGTCTCGAGGAGTCCCGAACAACCTTGACTAGCGCGCGGCGGCAAGACAT 1322  
QY 1032 GATCAATTG 1040  
DB 1323 TAAGACTAG 1331

RESULT 14  
US-09-883-096-6  
Sequence 6, Application US/09883096  
Patent No. US20020110883A1  
GENERAL INFORMATION:  
APPLICANT: Beraud, Christophe  
APPLICANT: Craven, Andrew

APPLICANT: Yu, Ming  
APPLICANT: Sakowicz, Roman  
APPLICANT: Patel, Umesh A.  
APPLICANT: Davies, Katherine A.  
TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE  
FILE REFERENCE: 020552-001410US  
CURRENT APPLICATION NUMBER: US/09/883,096  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 09/594,655  
PRIOR FILING DATE: 2000-06-15  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1152  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: HsKip3a  
OTHER INFORMATION: fragment  
OTHER INFORMATION: Nucleotide sequence of HsKip3a fragment used in  
OTHER INFORMATION: ATPase assay (Figure 5).  
US-09-883-096-6

Query Match 13.1%; Score 136.6; DB 9; Length 1152;  
Best Local Similarity 51.2%; Pred. No. 3.2e-33;  
Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;  
QY 196 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGTTTCAGTGCAG 255  
DB 214 TTTGACCGGGTCTTTGGCGAGCGGCCACCAACAGGACGTGTTCAGCAGCACACGCGAC 273  
QY 256 CCCATCCTAAGGCATCTGTGGAAGGGCAGATGCCAGTGTGTCTGCTATGACCCACACA 315  
DB 274 AGGCTCTGGACAGCTTCTCCAGGGCTACACTGTCTAGTGTTCAGTACGGGGCCACC 333  
QY 316 GAGCTGGGAAGACGACACAATGTCTGGGAGCCAGAGCAACCTGGGGTGTATCCCGCGG 375  
DB 334 GGGGCTGGGAAGACACACACCATGCTGGGAAGGGAGGGGACCCCGGCATCATGTA--- 389  
QY 376 GCTCTCATGGACCTCTCTGAGCTCAGAGGAGGAGGTGCCGAGGGCGGCCCATGGGCC 435  
DB 390 -----CCTACCCACCGTGGAACTGTATACGGCGCTGGAGCGCCGCGCAGCAGGAGAC 444  
QY 436 CTCTTCTGTCACCATGCTTTACCTAGAGATCTACAGGAGAGGTATTAGACCTCTCTGGAC 495  
DB 445 TTGAGGTCTCTATCAGCTACAGGAGGTGTATATGAACAGATCCATGACCTCTCTGGAG 504  
QY 496 CTGCTTCGGGAGACTGGTAAATCCGAGAAGACTGCCGGGGAAATCTCTGATTCGGGGT 555  
DB 505 C---CCAAGGGGCGCTTGCCTCCGAGGAGCCCGACAAAGGGGGTGGTGTGCAAGGA 561  
QY 556 CTCTCCAGAACCCCATCAGTAGCTTTGCTGATTTTGAGGGGCACTCTCTGCCAGCCAGT 615  
DB 562 CTCTTCTTCCACGAGCCAGCTCAGCCGAGCAGTCTGGAGATATGACACGAGGGGAAC 621  
QY 616 CGAAATTCGAGCTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGAGTATGCTGTG 675  
DB 622 CGTAACCGCAGCAGCAGCAGCCCACTGATGCAACGCGACTCTCTCCGCTCCCATGCCATC 681  
QY 676 CTCTCTGGTCAAGGTGAGCAGCGGGAACTTTTGGCCCCCATTTTCGCCAG-----CGAGAG 729  
DB 682 TTCCAGATCTTTGTGAAGCAGCAGGACCGGGTTCAGGAGTCCAGCCAGGCTGCTCCAGGTG 741  
QY 730 GGAATACTTACTGATTTGACTTGGCTGGGTGAGGACAAACCGGCGCAGCAGCAACAG 789  
DB 742 GCCAAGATGAGCTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 801  
QY 790 GGCCTTCGGCTAAAGAGAGAGTGGAGCCATCAACACCTCTCTGTTTCTCTGGGCAAGTG 849  
DB 802 GGGGAGCGCTCGGGAGGGGGGCGCAACATCAACCGCTCTCTGCTGGCGCTCATCAAGCTC 861  
QY 850 GTAGATGCGCTGAATCAGG-----GCCTCCCTCTGTTACCTTTATCGGGAACAGAG 900

Db 862 CTCATGCTTGGCGGATGCAAGGCGCGCAAGACCCATGTGCGCTTACCGGACAGCAA 921  
 QY 901 CTCACCTCGCTATTGACGACTCTCTGGGTGGCTCAGCCCAAGATATCCTTATTGCGCAAC 960  
 Db 922 CTGACCCGCGCTGCTCAAGACTCCCTCGGGGCAACTGCGGACAGTATGATCGTGCC 981  
 QY 961 ATTGCCCTGAGAGCGCTTCTACCTAGACACAGTCTCGGCACTCAACTTTCTGCGCAGG 1020  
 Db 982 ATCAGCCCTCCAGCTGACCTACGAGGACAGTACACACCTCAATATGCGGACCGG 1041  
 QY 1021 TCCAAGGAGGTGA 1033  
 Db 1042 GCCAAGGAGATCA 1054

RESULT 15

US-09-883-096-1  
 ; Sequence 1, Application US/09883096  
 ; Patent No. US20020110883A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Craven, Andrew  
 ; APPLICANT: Yu, Ming  
 ; APPLICANT: Sakowicz, Roman  
 ; APPLICANT: Patel, Umesh A.  
 ; APPLICANT: Davies, Katherine A.  
 ; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 020552-00141005  
 ; CURRENT APPLICATION NUMBER: US/09/883,096  
 ; CURRENT FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: US 09/594,655  
 ; PRIOR FILING DATE: 2000-06-15  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 4108  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Nucleic acid sequence of human kinesin motor  
 ; OTHER INFORMATION: protein gene HsKip3a (Figure 1).  
 ; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a gene.  
 US-09-883-096-1

Query Match 13.1%; Score 136.6; DB 9; Length 4108;  
 Best Local Similarity 51.2%; Pred. No. 4.7e-33;  
 Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;  
 QY 196 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCGAGTTTCAGTGCAG 255  
 Db 344 TTTGACCGGGTCTTTGGCGAGGCGGCCACCCCAACAGGACGTGTTCCAGCACACACGCAC 403  
 QY 256 CCATCCTTAAGCATTGTTGAAGGGGAGGAGTATGCGAGTGTCTTGTCTATGAGCCACA 315  
 Db 404 AGCGTCTGGAGCAGTTCCTCCAGGGCTACAACTGCTCAGTGTGTTGCTTACGGGCGCAC 463  
 QY 316 GGAGCTGGGAGACACACACATGCTGGGAGCGCCAGGACCACTGGGGTGTATCCCGGG 375  
 Db 464 GGGGCTGGGAGACACACACATGCTGGGAGGGAGGGGACCCCGGCATCATGTA---- 519  
 QY 376 GCTCTCATGGACCTCTGCGAGCTCAAGGGAGGAGGGTGGCGAGGGCGGCGCATGGGCC 435  
 Db 520 -----CCTGACCACCGTGAACCTGTACAGGCGCTGGAGGCCCGCCAGCAGGAGAGCAC 574  
 QY 436 CTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGGAGGATATTAGACTCTCTGAC 495  
 Db 575 TTGAGGTGCTCATCAGCTTACGAGAGGTGTATATGAACAGATCCATGACCTCTCTGGAG 634  
 QY 496 CCGTCTCGGAGACCTCGTATCGAGAGACTCGCGGGGGAATCTCTGATTCGGGT 555  
 Db 635 C---CAAAGGGGCCCCCTTGCCATCCCGAGGACCCCGACAGAGGGGGTGGTGTCAAGGA 691  
 QY 556 CTCTCCAGAGCCCATCAGTAGCTTTGTGATTTTGTAGCGGCACTTCTCTGCGACCCAGT 615

Db 682 CTTTCTTTCCACCAGCAGCCTCAGCCAGAGCAGCTGTGGAGATACTGACACAGGGGGAC 751  
 QY 616 CGAAATCGGACTGTAGAGCCACCGGCTCAACCAAGCGCTCCTCCGCGAGTCACTGTGTG 675  
 Db 752 CGTAACCGCAGCAGCAGCAGCAGCAGTGTGCGGCGCTTCTCTCGGCTCCCATGCCATC 811  
 QY 676 CTCCTGTGTCAGGTGAGCAGGGGAAAGTGTGGCCCCATTTTCGCCAG-----CGAGAG 729  
 Db 812 TTCCAGATCTTTGTGAAGCAGCAGGACCGGGTTCCAGGACTGACCCAGGCTGTTCAGGTG 871  
 QY 730 GGAAACCTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAACCGCGCAGCAGGCAACAG 789  
 Db 872 GCCAAGATGAGCTGATTGACCTGGCTGGCTCAGAGCGGGCATCCAGACACCCATCGCAG 931  
 QY 790 GGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTTGGGCAAGTG 849  
 Db 932 GGGGAGCGGCTGCGGGAGGGGCGCAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 991  
 QY 850 GTAGATGCGCTGAATCAGG-----GCCTCCCTCGTGTACCTTATCGGGACAGCAAG 900  
 Db 992 CTCAATGCTTTGGCGGATGCAAGGGCGCGCAAGACCCCATGTGCTTACCGGACAGCAA 1051  
 QY 901 CTCACCTCGCTATTGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCCTTATTGCCAAC 960  
 Db 1052 CTGACCCGCTGCTCAAGACTTCCCTCGGGGCAACTGCGGACAGTATGATGCTGCTGCC 1111  
 QY 961 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGG 1020  
 Db 1112 ATCAGCCCTCCAGCCTGACCTACGAGGACAGTACAAACACCCCTCAAAATATGCCGACCGG 1171  
 QY 1021 TCCAGGAGGTGA 1033  
 Db 1172 GCCAAGGAGATCA 1184

Search completed: November 11, 2004, 01:20:41  
 Job time : 558.617 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 19:47:28 ; Search time 6796.66 Seconds  
(without alignments)  
10701.074 Million cell updates/sec

Title: US-10-797-893-3  
Perfect score: 1538  
Sequence: 1 atgcagcgcggcggtctg.....accattgtcccaaatgtga 1538

Scoring table: IDENTITY\_NUC  
Gapop 10\_0, Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues  
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_phi.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_to.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1538	100.0	1538	6 AR210053	BT007259 Homo sapi
2	1530	99.5	1998	9 BT007259	BT007888 Synthetic
3	1530	99.5	1998	12 BT007888	AR304057 Sequence
4	1530	99.5	2097	6 AR304057	AB017430 Homo sapi
5	1530	99.5	2097	9 AB017430	BC028155 Homo sapi
6	1530	99.5	2117	9 BC028155	BC004352 Homo sapi
7	1530	99.5	2134	9 BC004352	AR210055 Sequence
8	1461	95.0	1464	6 AR210055	AR210052 Sequence
9	1112.4	72.3	1115	6 AR210052	BC003427 Mus muscu
10	1074.8	69.9	2086	10 BC003427	AR210054 Sequence
11	1035.4	67.3	1041	2 AR210054	AC101919 Mus muscu
c 12	885.2	57.6	151041	2 AC101919	AC101752 Mus muscu
c 13	885.2	57.6	164759	10 AC101752	AC101908 Mus muscu
c 14	744	48.4	297639	2 AC101908	AC102127 Mus muscu
15	706.8	46.0	196674	10 AC102127	BC063896 Xenopus t
16	511.6	33.3	2756	5 BC063896	AJ249841 Xenopus t
17	510.2	33.2	2583	5 XLA249841	BC043733 Xenopus 1
18	508.6	33.1	2565	5 BC043733	BC073177 Xenopus 1
19	508.6	33.1	2640	5 BC073177	

20	501	32.6	2593	5 AF267850	AF267850 Xenopus 1
21	501	32.6	2661	5 XLA249840	AJ249840 Xenopus 1
22	501	32.6	2781	5 BC070549	BC070549 Xenopus 1
23	501	32.6	2819	5 AF267849	AF267849 Xenopus 1
c 24	395	25.7	68916	2 AC101927	AC101927 Mus muscu
25	376.4	24.5	386	6 AX247007	AX247007 Sequence
c 26	365	23.7	174007	9 AC023831	AC023831 Homo sapi
c 27	365	23.7	174363	9 AC009133	AC009133 Homo sapi
28	361.8	23.5	3396	9 AB01733382	AB017334 Homo sapi
29	353	23.0	105485	9 HUAC002301	AC002301 Homo sapi
30	346.4	22.5	444	10 AF013119	AF013119 Mus muscu
31	265	17.2	228783	2 AC094377	AC094377 Rattus no
c 32	265	17.2	235806	2 AC133764	AC133764 Rattus no
c 33	265	17.2	289504	2 AC123480	AC123480 Rattus no
c 34	262.4	17.1	195987	10 AC122537	AC122537 Mus muscu
c 35	262.4	17.1	226601	10 AC122863	AC122863 Mus muscu
36	208.8	13.6	1982	3 AK114179	AK114179 Clona int
37	185.6	12.1	2095	6 AX833125	AX833125 Sequence
38	185.6	12.1	2095	9 AK094619	AK094619 Homo sapi
39	165	10.7	146415	5 AL929469	AL929469 Zebrafish
40	165	10.7	176906	2 CR354587	CR354587 Danio rer
41	153.8	10.0	3358	10 BC057614	BC057614 Mus muscu
42	145.8	9.5	435	5 MSU34658	U34658 Morone saxa
43	145.2	9.4	44004	2 AC092563	AC092563 Hylobates
44	143	9.3	183586	9 AC092562	AC092562 Papio ham
45	139	9.0	9701	9 AB01733383	AB017335 Homo sapi

ALIGNMENTS

RESULT 1  
AR210053  
LOCUS AR210053 1538 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 3 from patent US 6387644.  
ACCESSION AR210053  
VERSION AR210053.1 GI:21512183  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1538)  
AUTHORS Beraud, C.  
TITLE Motor proteins and methods for their use  
JOURNAL Patent: US 6387644-A 3 14-MAY-2002;  
FEATURES  
Location/Qualifiers  
source 1..1538  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match	100.0%	Score 1538;	DB 6;	Length 1538;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1538;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCAGCGCGCGGCTCGACGCGAGCGAGCGAGATGCGGCGAGCTTCAGCGG	60	
Db	1	ATGCAGCGCGCGGCTCGACGCGAGCGAGCGAGATGCGGCGAGCTTCAGCGG	60	
Qy	61	CGGCGATCTCAGGAGCTGGTGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCTGCTGCTCCAC	120	
Db	61	CGGCGATCTCAGGAGCTGGTGTGCTGTGGCTTAAGCAAGATTGGAGCTACTCTGCTGCTCCAC	120	
Qy	121	CTCAGCTCGCGTAAAGGTGGTGTGCGACTGCGGCCAATTTGTGATGGAACAGCGGGAG	180	
Db	121	CTCAGCTCGCGTAAAGGTGGTGTGCGACTGCGGCCAATTTGTGATGGAACAGCGGGAG	180	
Qy	181	CAAGTGATCCCCCTGTGTGCGGGGCGAGGACAGCTGCTCTAGAGATTGCTACTGGA	240	
Db	181	CAAGTGATCCCCCTGTGTGCGGGGCGAGGACAGCTGCTCTAGAGATTGCTACTGGA	240	
Qy	241	GGAACACACAGGAGACTCTCAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTC	300	

Db 241 GGAACACCAGGAGCTCTCAATACCAGTTGATGCTTCTATGGGAGGAGTACTC 300  
Qy 301 AGCAGGACATCTATGCAAGTTTCAGTGACAGCCATCTTAAGGCACTTCTGTAAGGGGCGA 360  
Db 301 AGCAGGACATCTATGCAAGTTTCAGTGACAGCCATCTTAAGGCACTTCTGTAAGGGGCGA 360  
Qy 361 ATGCCAGTGTCTTCTATGACCCACAGAGCTGGGAGAGCCACACATGCTGGGCA 420  
Db 361 ATGCCAGTGTCTTCTATGACCCACAGAGCTGGGAGAGCCACACATGCTGGGCA 420  
Qy 421 GCCCAGAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAACAAGG 480  
Db 421 GCCCAGAGCAACCTGGGGTGATCCCGGGGCTCTCATGGACCTCTCGAGCTCAACAAGG 480  
Qy 481 AGGAGGTGCGAGGGCGGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGATCT 540  
Db 481 AGGAGGTGCGAGGGCGGCGCATGGGCGCTTCTGTCAACATGCTTACCTAGATCT 540  
Qy 541 ACCAGGAGAGGTATTAGACCTCTGACACCTGCTTGGGAGACCTGTAATCCGAGAAG 600  
Db 541 ACCAGGAGAGGTATTAGACCTCTGACACCTGCTTGGGAGACCTGTAATCCGAGAAG 600  
Qy 601 ACTCGCGGGGAATATCTGATCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660  
Db 601 ACTCGCGGGGAATATCTGATCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTG 660  
Qy 661 ATTTTGAAGGCACTTCTGACGACAGTGAATTCGGACTGTAGAGCCACCCGGCTCA 720  
Db 661 ATTTTGAAGGCACTTCTGACGACAGTGAATTCGGACTGTAGAGCCACCCGGCTCA 720  
Qy 721 ACCAGCGCTCTCCCGGAGTATCTGCTGCTCTGCTCAAGTGTGACCGGGAAGCTT 780  
Db 721 ACCAGCGCTCTCCCGGAGTATCTGCTGCTCTGCTCAAGTGTGACCGGGAAGCTT 780  
Qy 781 TGGCCCCATTTGCGCAGGAGGAGAACTCTACCTGATTGACTTGGCTGGGTTCAGAGG 840  
Db 781 TGGCCCCATTTGCGCAGGAGGAGAACTCTACCTGATTGACTTGGCTGGGTTCAGAGG 840  
Qy 841 ACAACCGGCGCAGGCAACAGGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACCT 900  
Db 841 ACAACCGGCGCAGGCAACAGGGGCTTGGCTTAAAGAGAGTGGAGCCATCAACCT 900  
Qy 901 CCTGTTTGTCTCGGCAAGTGTAGATGGCTGGAATCAGGGCTCCCTCGTGTACTT 960  
Db 901 CCTGTTTGTCTCGGCAAGTGTAGATGGCTGGAATCAGGGCTCCCTCGTGTACTT 960  
Qy 961 ATCGGACAGCAAGCTCACTGCGCTATTGACGAGTCTCTGGGTGGCTCAGCCACAGTA 1020  
Db 961 ATCGGACAGCAAGCTCACTGCGCTATTGACGAGTCTCTGGGTGGCTCAGCCACAGTA 1020  
Qy 1021 TCCTTATGCGCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
Db 1021 TCCTTATGCGCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
Qy 1081 ACTTGTGTCAGGTCAGAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTCGAGC 1140  
Db 1081 ACTTGTGTCAGGTCAGAGAGTGTATCAATCGGCTTTTACCAATGAGAGCTCGAGC 1140  
Qy 1141 CTCATGCTTTGGGACCTTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCAA 1200  
Db 1141 CTCATGCTTTGGGACCTTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGCAA 1200  
Qy 1201 AGAGAGCCGAGGCGCTGAGGAGAGAGATTTGGAGCCCTGAGCCCATGCGAGCTCCAG 1260  
Db 1201 AGAGAGCCGAGGCGCTGAGGAGAGAGATTTGGAGCCCTGAGCCCATGCGAGCTCCAG 1260  
Qy 1261 CCTGTGCTCCAGAACTCAGCCCTTACAGAGAGTAAAGCAGCATGACGCCCGCCATGC 1320  
Db 1261 CCTGTGCTCCAGAACTCAGCCCTTACAGAGAGTAAAGCAGCATGACGCCCGCCATGC 1320  
Qy 1321 TGGAGCGCTCTCAGCTTGGACCTGTGCTTGGCTCCCGAGGAGCAGCGGGGCCCTC 1380  
Db 1321 TGGAGCGCTCTCAGCTTGGACCTGTGCTTGGCTCCCGAGGAGCAGCGGGGCCCTC 1380

Qy 1381 TGTTCAGTACCCAAAGCAGAGCGGATGCTGCTAATGAAGACAGTAAAGAGGACC 1440  
Db 1381 TGTTCAGTACCCAAAGCAGAGCGGATGCTGCTAATGAAGACAGTAAAGAGGACC 1440  
Qy 1441 TAGAGATTGAGAGCTTAAAGCAAGCAAAAGAACTGGAGGCGCAAGATGTTGCCCGAGA 1500  
Db 1441 TAGAGATTGAGAGCTTAAAGCAAGCAAAAGAACTGGAGGCGCAAGATGTTGCCCGAGA 1500  
Qy 1501 AGCTGAGGAAAAGGAGAACCATTTGTCCTCCCAATGTGA 1538  
Db 1501 AGCTGAGGAAAAGGAGAACCATTTGTCCTCCCAATGTGA 1538

## RESULT 2

BT007259

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

1998 bp mRNA linear PRI 13-MAY-2003  
Homo sapiens kinesin-like 4 mRNA, complete cds.

BT007259 Homo sapiens kinesin-like 4 mRNA, complete cds.

BT007259.1 GI:305933356

FLI\_CDNA.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1998)

Kalnina, N., Chen, X., Rolfes, A., Halleck, A., Hines, L., Eisenstein, S.,

Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y.,

Phelan, M., and Farmer, A.

Cloning of human full-length CDSs in BD Creator(TM) System Donor

vector

Unpublished

2 (bases 1 to 1998)

Kalnina, N., Chen, X., Rolfes, A., Halleck, A., Hines, L., Eisenstein, S.,

Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y.,

Phelan, M., and Farmer, A.

Direct Submission

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow

Circle, Palo Alto, CA 94303, USA

This CDS clone is a part of a collection of human full length

expression clones generated by BD Biosciences Clontech and the

Harvard Institute of Proteomics. Each CDS has been cloned in two

forms: with and without stop-codon (to allow fusion with C-terminal

tag). The CDS has been directionally cloned using BD In-Fusion(TM)

cloning system between the SalI and HindIII sites of the pDNR-DUAL

vector. Additional sequences in the clone: 'ACC' after SalI site

and before 'ATG' to provide Kozak consensus sequence; 'GG' after

last codon and before HindIII site to maintain reading frame.

Clone distribution: <http://bioinfo.clontech.com/orfclones>.

Location/Qualifiers

1..1998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="BD Creator(TM) CDS Library derived from MGC

collection"

/lab\_host="DH5alpha T1 resistant"

/notes="Vector: pDNR-Dual"

1..1998

/codon\_start=1

/product="kinesin-like 4"

/protein\_id="AAP35923.1"

/db\_xref="GI:30593357"

/translation="MAAGSTQQRREMAAASAAI SGAGRCRLSKITGATRPPPARV

RVAVRPFVDGTAGADPPCVGRMDSCSLEIANWRNHQETLYQDFAGTGERSTOOD

IVAGSQPILRHLEQNASVLAAYPTGTMLGSPQGVIPALMDLLQLTRE

EGAGRPWALVTVMLEIYQEVKLLDLPDAGSDLVIREDCRGNILIPGLSOKPISF

ADPERHFLPASRNTVGATFLNQRSSRHVLLVKVDORERLAPPRORCKLYLIDIA

GSNDRITNGKRLKESGINTSLFVLGKVDALNQGLPRVPRRSLKRLQLQDSIG

GSASHILIANIAPERFYLDTVSNLFPAAKSKVINRPFNTNSLQHPALGPKVLKSLG

LLGPPAKARGPEEEI GSPPEMPAASOKLSPLQKLSMDPAMLEKLLSLDRLL

ASQSGQAPILLTPKRRMVLMTVBKDLIEIRLTKQKELEAKMLAQAEKENHC



last codon and before HindIII site to maintain reading frame.  
Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES  
source

```

1..1998
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH00245L1.0"
/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1..>1998
/note="Mutations: 1997:Stop->Leu"
/codon_start=1
/transl_table=11
/product="Homo sapiens kinesin-like 4"
/protein_id="AAP36560.1"
/db_xref="GI:30584615"
/translation="MAAGSTQQRREMAAASAAISGACRCLSKIGATRRPPARV
RVAVRLPFDVDTAGADPPCVRGMSCSLEIANWRNHQTLKTFQDFYGERSTQD
IYAGVOPILRLHLEGQNASVLYAGPTGAKTHMLGSPQPGVIPALMDLQLQFRE
EGAEGRWALSVTMSYLEIYQEKVLDLDPASGLVIREDCRGNILIPGLSQPISSF
ADFERHLPASRLNRTVGATRLNRSRSHAVLLVKVDQRLAPFRQREKGLYLIDLA
GSDENRTGNKGLRKESGALNTSLFVLGVLDALNOGLRPVPROSKLTRLQDSIG
GSAHILIANAPERRYLDTVSALNFAARSKEVINRPTNESLOPHALGPVKLSQKE
LLGPEAKRAGPEEESIGSPFEMPAASASOKLSPLQKUSSMDPAMLAQBAEKENHC
ASQSGAPLLSTPKRMRIMTKTVEEKLEIEIRLTKQKELEAKMLQAEKBNHC
PTMLRPLSHRTVTGAKPLKKAIVMPLQLIQEASPNAEIHLINKRKRKLSLDAL
EPBEKARDCHWELQISPELLAHGROKILDLNENSGARDLSLRIGPKKAQLIVGWREL
HGPPSQVEDLERVEGITGKQWESFLKANILGLAQRCSASL"

```

ORIGIN

```

Query Match      99.5%; Score 1530; DB 12; Length 1998;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GCCCGGGCGCTCGACGACGACGAGCGCGGAGATGCGGAGCTTCCAGCGGGCGG 65
DB 4 GCCCGGGCGCTCGACGACGACGAGCGCGGAGATGCGGAGCTTCCAGCGGGCGG 63

QY 66 ATCTCAGGAGTGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCA 125
DB 64 ATCTCAGGAGTGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTCCACCTCCA 123

QY 126 GCTCGCTAAGGTGGCTGTGCGACTCGGCGCATTTGTGATGGAACAGCGGGAGCAAGT 185
DB 124 GCTCGCTAAGGTGGCTGTGCGACTCGGCGCATTTGTGATGGAACAGCGGGAGCAAGT 183

QY 186 GATCCCGCTGTGTGCGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAAC 245
DB 184 GATCCCGCTGTGTGCGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGAGGAAC 243

QY 246 CACAGAGACTCTCAAAATACAGTTTGTATGCGCTTCTATGGGAGAGAGTACTCAGCAG 305
DB 244 CACAGAGACTCTCAAAATACAGTTTGTATGCGCTTCTATGGGAGAGAGTACTCAGCAG 303

QY 306 GACATCTATCAGGTTTCAAGTGCAGCCCATCTTAAGGCATTTGCTGNAAGGCGAATGCC 365
DB 304 GACATCTATCAGGTTTCAAGTGCAGCCCATCTTAAGGCATTTGCTGNAAGGCGAATGCC 363

QY 366 AGTGTGCTTGCCTATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCCA 425
DB 364 AGTGTGCTTGCCTATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCCA 423

QY 426 GAGCAACTGGGTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAACAGGGAGGAG 485
DB 424 GAGCAACTGGGTGATCCCGGGGCTCTCATGGAACCTCTGAGCTCAACAGGGAGGAG 483

QY 486 GGTGCGGAGGCGGCGCATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACCAG 545
DB 484 GGTGCGGAGGCGGCGCATGGGCGCTTCTGTCCACCATGCTTACCTAGAGATCTACCAG 543

```

RESULT 4  
AR304057  
LOCUS  
DEFINITION

Sequence 34 from patent US 6544766.  
2097 bp  
DNA  
linear  
PAT 12-JUN-2003

```

546 GAGNAGTATTAGACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGC 605
544 GAGAAGGTATTAGACCTCTCGGACCTGCTTCGGGAGACCTGGTAATCCGAGAAGACTGC 603
606 CGGGGGAATATCTCTGATTTCGGGTCTCTCCAGAAAGCCCATCAGTACGCTTTGCTGATTTT 665
604 CGGGGGAATATCTCTGATTTCGGGTCTCTCCAGAAAGCCCATCAGTACGCTTTGCTGATTTT 663
666 GAGCGGCACTTCTTCGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAG 725
664 GAGCGGCACTTCTTCGCGAGCCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAG 723
726 CGCTCTCTCCGCGAGTCATGCTGTGCTCTCAAGTGAACCGGGAACGTTTGGCC 785
724 CGCTCTCTCCGCGAGTCATGCTGTGCTCTCAAGTGAACCGGGAACGTTTGGCC 783
786 CCATTTTCGCGAGGAGGAAATCTTACCTGATTCGCTTGGCTGGGTGAGAGCAAC 845
784 CCATTTTCGCGAGGAGGAAATCTTACCTGATTCGCTTGGCTGGGTGAGAGCAAC 843
846 CGCGGCAAGCAAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACTCCCTG 905
844 CGCGGCAAGCAAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACTCCCTG 903
906 TTTGTCTCGGCAAGTGTGATGCGCTGAATCAGGGCCCTCCCTGCTGACCTTATCGG 965
904 TTTGTCTCGGCAAGTGTGATGCGCTGAATCAGGGCCCTCCCTGCTGACCTTATCGG 963
966 GACAGCAAGTCTACTCGCTTATTCGAGGACTCTCTGGTGGCTCAGGCCACAGTATCTT 1025
964 GACAGCAAGTCTACTCGCTTATTCGAGGACTCTCTGGTGGCTCAGGCCACAGTATCTT 1023
1026 ATTGCAACATTCGCTCAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTT 1085
1024 ATTGCAACATTCGCTCAGAGACGCTTCTACTAGACACAGTCTCCGCACTCAACTTT 1083
1086 GCTGCCAGGTCGAAGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTCGAGCTCAT 1145
1084 GCTGCCAGGTCGAAGAGGTGATCAATCGGCCCTTTTACCAATGAGAGCTCGAGCTCAT 1143
1146 GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGCAAGAGA 1205
1144 GCCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACAGAGCAAGAGA 1203
1206 GCCGAGGCTCAGAGAGAGAGATGGAGCCCTTGGAGCCCTTGGAGCTCAGGCTCT 1265
1204 GCCGAGGCTCAGAGAGAGAGATGGAGCCCTTGGAGCCCTTGGAGCTCAGGCTCT 1263
1266 GCCTCCAGAACTCAGCCCTTACAGAGCTAAGCAGATGACCCGCGCATGCTGGAG 1325
1264 GCCTCCAGAACTCAGCCCTTACAGAGCTAAGCAGATGACCCGCGCATGCTGGAG 1323
1326 CGCTCTCTCAGCTTGGACCTGCTTGGCTTCCAGGGAGCCAGGGGGGCCCCCTCTGTTG 1385
1324 CGCTCTCTCAGCTTGGACCTGCTTGGCTTCCAGGGAGCCAGGGGGGCCCCCTCTGTTG 1383
1386 AGTACCCCAAGCAGAGCGATGCTGCTTAATGAACACAGTGAAGAGAGGAGCTTAGAG 1445
1384 AGTACCCCAAGCAGAGCGATGCTGCTTAATGAACACAGTGAAGAGAGGAGCTTAGAG 1443
1446 ATTGAGAGGCTTAAGACGAGGAAAGAAATGAGAGCCCAAGATGTTGGCCCAAGAGGCT 1505
1444 ATTGAGAGGCTTAAGACGAGGAAAGAAATGAGAGCCCAAGATGTTGGCCCAAGAGGCT 1503
1506 GAGGAAAGGAGAACCAATTGTCCCAATG 1535
1504 GAGGAAAGGAGAACCAATTGTCCCAATG 1533

```











Db 489 GGTGCCGAGGCGCGCCATCGGCCCTTTCTGTCAACATGCTTTACCTAGAGATCTACCCAG 548  
QY 546 GAGAGGTATTAGACCTCTCGGACCTCTGCTTCGGAGACCTGTTAATCCGAGAGACTGC 605  
Db 549 GAGAAGGTATTAGACCTCTCGGACCTCTGCTTCGGAGACCTGTTAATCCGAGAGACTGC 608  
QY 606 CGGGGGAATATCTGATTCTCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTT 665  
Db 609 CGGGGGAATATCTGATTCTCGGGTCTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTT 668  
QY 666 GAGCGGCACCTCTCGCGACCGCAGTCGAAATCGGACTGTAGGACCCACCGGCTCAACCAG 725  
Db 669 GAGCGGCACCTCTCGCGACCGCAGTCGAAATCGGACTGTAGGACCCACCGGCTCAACCAG 728  
QY 726 CGCTCCTCCGCGAGTCATGCTGCTCTGCTCAAGTGCACAGCGGGAACGTTTGGCC 785  
Db 729 CGCTCCTCCGCGAGTCATGCTGCTCTGCTCAAGTGCACAGCGGGAACGTTTGGCC 788  
QY 786 CCATTTGCGCAGCGAGGAGGAAACTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAC 845  
Db 789 CCATTTGCGCAGCGAGGAGGAAACTCTACCTGATTGACTTGGCTGGGTGAGAGGACAAC 848  
QY 846 CGGCGCACAGGCAACAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACCTCCCTG 905  
Db 849 CGGCGCACAGGCAACAGGCGCTTCGGCTAAAAGAGAGTGGAGCCATCAACCTCCCTG 908  
QY 906 TTTGTCTCGGGAAGTGTAGATCGCTGAATCAGGGCCCTCCCTGCTGTGACCTTATCGG 965  
Db 909 TTTGTCTCGGGAAGTGTAGATCGCTGAATCAGGGCCCTCCCTGCTGTGACCTTATCGG 968  
QY 966 GACAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGGCCACAGTATCCCT 1025  
Db 969 GACAGCAAGCTCACTCGCTTATTCAGGAGCTCTCTGGTGGCTCAGGCCACAGTATCCCT 1028  
QY 1026 ATTGCCAACATTCGCCCTGAGAGACGCTTCTACCTAGACACAGCTCTCCGACCTCAACTT 1085  
Db 1029 ATTGCCAACATTCGCCCTGAGAGAGGCTTCTACCTAGACACAGCTCTCCGACCTCAACTT 1088  
QY 1086 GCTGCCAGGTCGAAGAGGTGATCAATCGCCCTTTTACCAATGAGAGCCTGACGCTCAT 1145  
Db 1089 GCTGCCAGGTCGAAGAGGTGATCAATCGCCCTTTTACCAATGAGAGCCTGACGCTCAT 1148  
QY 1146 GCCTTGGACCTGTTAAGCTGCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGA 1205  
Db 1149 GCCTTGGACCTGTTAAGCTGCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGA 1208  
QY 1206 GCCCGAGGCCCTGAGGAAGAGGAGATTGGAGGCCCTGAGCCCATGGCAGCTCCAGCCTCT 1265  
Db 1209 GCCCGAGGCCCTGAGGAAGAGGAGATTGGAGGCCCTGAGCCCATGGCAGCTCCAGCCTCT 1268  
QY 1266 GCCTCCGAGAACTCAGCCCTTACAGAGCTAAGCGCATGACACCGGCCATGCTGGAG 1325  
Db 1269 GCCTCCGAGAACTCAGCCCTTACAGAGCTAAGCGCATGACACCGGCCATGCTGGAG 1328  
QY 1326 CGCTCCTCAGCTTGGACCTGCTTGGCTCCAGGGGAGCAGGGGGCCCTCTGTTG 1385  
Db 1329 CGCTCCTCAGCTTGGACCTGCTTGGCTCCAGGGGAGCAGGGGGCCCTCTGTTG 1388  
QY 1386 AGTACCCCAAGGAGAGCGGATGGTCTAATGAAGACAGTAGAAGAGAGGACCTAGAG 1445  
Db 1389 AGTACCCCAAGGAGAGCGGATGGTCTAATGAAGACAGTAGAAGAGAGGACCTAGAG 1448  
QY 1446 ATTGAGAGGCTTAAGACGAAGCAAAAGAACTGGAGCCCAAGATGTTGGCCCAAGAGGCT 1505  
Db 1449 ATTGAGAGGCTTAAGACGAAGCAAAAGAACTGGAGCCCAAGATGTTGGCCCAAGAGGCT 1508  
QY 1506 GAGGAAAAAGGAGAACCATTTGTCACCAATG 1535  
Db 1509 GAGGAAAAAGGAGAACCATTTGTCACCAATG 1538

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BC004352 2134 bp mRNA linear PRI 29-JUN-2004  
Homo sapiens kinesin family member 22, mRNA (cdna clone MGC:1573  
IMAGE:3535435), complete cds.  
BC004352  
BC004352.1 GI:13279307  
MGC.

REFERENCE  
AUTHORS

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2134)  
Krausberg, R.L., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,  
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

## TITLE

PROC. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932

JOURNAL  
PUBMED

2 (bases 1 to 2134)  
Strausberg, R.  
Direct Submission  
Submitted (01-MAR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth  
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,  
Tim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,  
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,  
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,  
Duane Smalusz, Jeff Stott, Miranda Tsai, George Yang, Jacquie  
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 8 Row: f Column: 1  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 6453817.

FEATURES  
source

1..2134  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:1573 IMAGE:3535435"  
/tissue\_type="lung, small cell carcinoma"  
/clone\_lib="NIH MGC.7"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"

gene	1. .2134	
	/gene="KIF22"	
	/note="synonyms: KID, OBP, OBP-1, OBP-2, KNSL4"	
	/db_xref="LocusID:3835"	
	/db_xref="MIM:603213"	
CDS	20. .2017	
	/gene="KIF22"	
	/codon_start=1	
	/product="kinesin family member 22"	
	/protein_id="AAH04352.1"	
	/db_xref="GI:13279308"	
	/db_xref="LocusID:3835"	
	/db_xref="MIM:603213"	
	/translation="MAAGSTQRRREMAAASAAISGAGRCRLSKIGATRRPPARV RYVRLRFVDGTAGASPPCVRGMDSCSLEIANRNHOETLKYOFDAFYGRSTOOD IYAGSVQPTLRHLLQONASVLAYGTGAGKTHWLGSPQPGVTPRALMDLLOITRE EGAGRPWALSVTMSYLEIOEKVLDLDPASGDLVIREDCRGNLILPGLSQKPISSP ADPERFLPASNRVTGATRLNQSRSHAVLLVKVQRERLAPFRQREGKYLIDLA GSEDRRTGNGLRUKESGAINTSIFVLGVKVDALNQGRLPVFPYRDSKLTRLQDLSL GSAHSILIANIAPERRFYLDVTALNFAARSKEVINRPTNESLQPHALGPVKLSQKE LLGPPKARAGPEEEIGSPPEMAAPASOKLSPLQKLSMDPAMLERLILSLDRLL ASQSGQAPLSTPKRERMLVMKTVEEKDLEIRLTKQKLEAKMLAQKAEKENHC PTMLRPLSHRTVTGAKPLKKAIVVPLQLIOEQAAAPNAEIIHLKNKGRKRLKSLDAL PPEEKAEDCWELQISPELLAHGRQKILDLNLEGSARDLSLQRIQPKQAQLIVGWREL HGFFSQVEDLERVEGITQKQMESFLKANILGAGRCGAS"	
ORIGIN		
Query Match	99.5%; Score 1530; DB 9; Length 2134;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1530; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	6	GCCGCGGGCGGCTCGACGAGGAGGCGGCGGAGATGGCGGCGAGCTTCAGCGCGGCG 65
DB	23	GCCGCGGGCGGCTCGACGAGGAGGCGGCGGAGATGGCGGCGAGCTTCAGCGCGGCG 82
QY	66	ATCTCAGAGAGCTGCTGCTGCGCTAAGCAAGATTGAGCTACTCTGCTGCTCACTCCA 125
DB	83	ATCTCAGAGAGCTGCTGCTGCTGCGCTAAGCAAGATTGAGCTACTCTGCTGCTCACTCCA 142
QY	126	GCTCCGCTAAGGGTGGCTGTGGAGCTGCGGCCATTTGGTGGAGCAAGCGGAGCAAGT 185
DB	143	GCTCCGCTAAGGGTGGCTGTGGAGCTGCGGCCATTTGGTGGAGCAAGCGGAGCAAGT 202
QY	186	GATCCCCCTGTGTGCGGGGCAATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 245
DB	203	GATCCCCCTGTGTGCGGGGCAATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 262
QY	246	CACAGGAGACTCTCAATATACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAG 305
DB	263	CACAGGAGACTCTCAATATACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAG 322
QY	306	GACATCTATCAGGTTTCAGTGAGCCCATCTTAAAGGCACTTCTGCTGGAAGGCGAGATGCC 365
DB	323	GACATCTATCAGGTTTCAGTGAGCCCATCTTAAAGGCACTTCTGCTGGAAGGCGAGATGCC 382
QY	366	AGTGTGCTTGCTATGAGCCCAACAGGAGCTGGGAGAGCGACACAATGCTGGGAGGCCCA 425
DB	383	AGTGTGCTTGCTATGAGCCCAACAGGAGCTGGGAGAGCGACACAATGCTGGGAGGCCCA 442
QY	426	GAGCAACTCTGGGTGATCCCGGGGCTCTCATGGACCTCTGCTGAGCTCACAAGGGAGAG 485
DB	443	GAGCAACTCTGGGTGATCCCGGGGCTCTCATGGACCTCTGCTGAGCTCACAAGGGAGAG 502
QY	486	GCTGCCGAGGCGCGGCTATGGCCCTTCTGTACCATGCTTCTAGAGATCTTACCAG 545
DB	503	GCTGCCGAGGCGCGGCTATGGCCCTTCTGTACCATGCTTCTAGAGATCTTACCAG 562
QY	546	GAGAGGTTATTAGACTCTCTGAGCCCTGCTCGGAGACCTGGTAACTCCGAGAGACTGCG 605
DB	563	GAGAGGTTATTAGACTCTCTGAGCCCTGCTCGGAGACCTGGTAACTCCGAGAGACTGCG 622
QY	606	CGGGGGAATATCTGATTCCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTT 665

DB	623	CGGGGGAATATCTGATTCCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTTT 682
QY	666	GAGCGGCACTTCTCGCAGCAGCTCGAATCGGATCTGTAGGAGCCACCCGGCTCAACCCAG 725
DB	683	GAGCGGCACTTCTCGCAGCAGCTCGAATCGGATCTGTAGGAGCCACCCGGCTCAACCCAG 742
QY	726	CGCTCTCTCCGCACTCATGCTGCTCTCTGCTCAAGGTTGAGCCAGCGGGAAGCTTTGGGC 785
DB	743	CGCTCTCTCCGCACTCATGCTGCTCTCTGCTCAAGGTTGAGCCAGCGGGAAGCTTTGGGC 802
QY	786	CAATTTCCGCACTCATGCTGCTCTCTGCTCAAGGTTGAGCTTTGGGTGAGGAGCAAC 845
DB	803	CAATTTCCGCACTCATGCTGCTCTCTGCTCAAGGTTGAGCTTTGGGTGAGGAGCAAC 862
QY	846	CGGCGCACAGGCAACAAAGGCGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 905
DB	863	CGGCGCACAGGCAACAAAGGCGCTTCGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTG 922
QY	906	TTTGTCTCGGCAAGTGGTAGATCGCTGCTCAAGGCGCTCCCTCGTGTACTCTTATCGG 965
DB	923	TTTGTCTCGGCAAGTGGTAGATCGCTGCTCAAGGCGCTCCCTCGTGTACTCTTATCGG 982
QY	966	GACAGCAAGCTCATCTCGCTTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGATTCCTT 1025
DB	983	GACAGCAAGCTCATCTCGCTTATTGAGGACTCTCTGGGTGGCTCAGCCCAAGATTCCTT 1042
QY	1026	ATTGCCAACATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1085
DB	1043	ATTGCCAACATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT 1102
QY	1086	GCTGCAGGCTCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCGAGGCTCAT 1145
DB	1103	GCTGCAGGCTCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTCGAGGCTCAT 1162
QY	1146	GCTTTGGGACCTTTAAGCTGTCTCAGAAAAGATTGTTGGTCCACAGAGCAAGAGA 1205
DB	1163	GCTTTGGGACCTTTAAGCTGTCTCAGAAAAGATTGTTGGTCCACAGAGCAAGAGA 1222
QY	1206	GCCGAGGCGCTGAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGAGCTCCAGGCTCT 1265
DB	1223	GCCGAGGCGCTGAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGAGCTCCAGGCTCT 1282
QY	1266	GCCTCCCAAGAACTCAGCCCTTACAGAAAGCTAAGCAGCATGGAGCCGCGCATGCTGGAG 1325
DB	1283	GCCTCCCAAGAACTCAGCCCTTACAGAAAGCTAAGCAGCATGGAGCCGCGCATGCTGGAG 1342
QY	1326	CGCTCTCTCAGCTTGGACCGTCTGCTTCCAGGGGAGCCAGGGGGCCCTCTGTTG 1385
DB	1343	CGCTCTCTCAGCTTGGACCGTCTGCTTCCAGGGGAGCCAGGGGGCCCTCTGTTG 1402
QY	1386	AGTACCCCAAGCGAGGAGCGGATGGTCTAATGAGCAGTAGAGAGAGAGGAGCTAGAG 1445
DB	1403	AGTACCCCAAGCGAGGAGCGGATGGTCTAATGAGCAGTAGAGAGAGAGGAGCTAGAG 1462
QY	1446	ATTGAGAGGCTTTAAGACGAAAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCCAAGGCT 1505
DB	1463	ATTGAGAGGCTTTAAGACGAAAGCAAAAGAACTGGAGGCGCAAGATGTTGGCCCAAGGCT 1522
QY	1506	GAGGAAAGGAGAACCATTTGTCACCAATG 1535
DB	1523	GAGGAAAGGAGAACCATTTGTCACCAATG 1552

RESULT 8				
LOCUS	AR210055	1464 bp	DNA	linear
DEFINITION	Sequence 7 from patent US 6387644.			
ACCESSION	AR210055			
VERSION	AR210055.1	GI:21512186		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			

REFERENCE	1 (bases 1 to 1464)	
AUTHORS	Beraud,C.	
TITLE	Motor proteins and methods for their use	
JOURNAL	Patent: US 6387644-A 7 14-MAY-2002;	
FEATURES	Location/Qualifiers	
source	1..1464	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	95.0%; Score 1461; DB 6; Length 1464;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1461; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	78 GGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCCAGCTCCAGCTCGCGTAAGG	137
DB	4 GGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCCAGCTCCAGCTCGCGTAAGG	63
QY	138 GTGGCTGTGGCTAAGCAAGATTGGATGGACAGCGGAGCAAGTGTCCCGCTGT	197
DB	64 GTGGCTGTGGCTAAGCAAGATTGGATGGACAGCGGAGCAAGTGTCCCGCTGT	123
QY	198 GTGGGGGCGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	257
DB	124 GTGGGGGCGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT	183
QY	258 CTCAATATACAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
DB	184 CTCAATATACAGTTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
QY	318 GGTTCAGTGCAGCCATCTTAAGCACTTCTGGAAGGGCAGATGCCAGTGTGCTTGC	377
DB	244 GGTTCAGTGCAGCCATCTTAAGCACTTCTGGAAGGGCAGATGCCAGTGTGCTTGC	303
QY	378 TATGGAACCAAGAGAGCTGGGAAGACGACACATATGCTGGGAGCCACCACTGGG	437
DB	304 TATGGAACCAAGAGAGCTGGGAAGACGACACATATGCTGGGAGCCACCACTGGG	363
QY	438 GTGATCCGGGGGCTCTCATGAGACTCTGAGAGCTCAAGGAGAGAGGTTGCCAGGGC	497
DB	364 GTGATCCGGGGGCTCTCATGAGACTCTGAGAGCTCAAGGAGAGAGGTTGCCAGGGC	423
QY	498 CGGCCATGGGCCCTTCTGTACCATCTCTTACCTAGAGATCTACAGGAGAGGTATTA	557
DB	424 CGGCCATGGGCCCTTCTGTACCATCTCTTACCTAGAGATCTACAGGAGAGGTATTA	483
QY	558 GACCTCTGGACCTTCTCGGAGACTGCTGTAATCCGAGAAAGACTCCCGGGGAATATC	617
DB	484 GACCTCTGGACCTTCTCGGAGACTGCTGTAATCCGAGAAAGACTCCCGGGGAATATC	543
QY	618 CTGATTCGGGTCTCTCCAGAACCCCATCAGTACGTTTGTGATTTTGAGCGGCATTC	677
DB	544 CTGATTCGGGTCTCTCCAGAACCCCATCAGTACGTTTGTGATTTTGAGCGGCATTC	603
QY	678 CTGCGACCGAGTCAAAATCCGAGCTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	737
DB	604 CTGCGACCGAGTCAAAATCCGAGCTAGGAGCCACCGGCTCAACAGCGCTCTCCCGC	663
QY	738 AGTCATGCTGTGCTCTCGGTCAAGGTGGACAGCGGGAAGCTTTGGCCCATTTTCGCCAG	797
DB	664 AGTCATGCTGTGCTCTCGGTCAAGGTGGACAGCGGGAAGCTTTGGCCCATTTTCGCCAG	723
QY	798 CGAGAGGGAATCTCTACCTGATTTGACTTTGGCTGGGTCAAGAGCAACCGCGGCACAGGC	857
DB	724 CGAGAGGGAATCTCTACCTGATTTGACTTTGGCTGGGTCAAGAGCAACCGCGGCACAGGC	783
QY	858 AACAGGCGCTTCGCGTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCGGC	917
DB	784 AACAGGCGCTTCGCGTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTCGGC	843
QY	918 AAAGTGTGTAGATCGGCTGAATCAGGGGCTCCCTCGTGTACCTTATCGGGAACAGCAAGCTC	977
DB	844 AAAGTGTGTAGATCGGCTGAATCAGGGGCTCCCTCGTGTACCTTATCGGGAACAGCAAGCTC	903
QY	978 ACTCGCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTATCTCTATTGCCACATT	1037
DB	904 ACTCGCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTATCTCTATTGCCACATT	963
QY	1038 GCCCTGTGAGAGCGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTTGTGTCAGGTCC	1097
DB	964 GCCCTGTGAGAGCGCTTCTTACCTAGACACAGTCTCCGCACTCAACTTTTGTGTCAGGTCC	1023
QY	1098 AAGGAGTGATCAATCGGCTTTTACCAATGAGAGCTCCAGCCTCATGCTTTGGAGCCT	1157
DB	1024 AAGGAGTGATCAATCGGCTTTTACCAATGAGAGCTCCAGCCTCATGCTTTGGAGCCT	1083
QY	1158 GTTAACTGTCTCAGAAAGAAATTTGCTTGCACAGAGCAAGAGAGCCGAGGCCCT	1217
DB	1084 GTTAACTGTCTCAGAAAGAAATTTGCTTGCACAGAGCAAGAGAGCCGAGGCCCT	1143
QY	1218 GAGGAAGAGGAGATTGGGAGCCTGAGCCATCGGAGCTCCAGCCTCTGCTCCAGAAA	1277
DB	1144 GAGGAAGAGGAGATTGGGAGCCTGAGCCATCGGAGCTCCAGCCTCTGCTCCAGAAA	1203
QY	1278 CTGAGCCCCCTACAGAAAGCTTAAGCATGGACCCCGGCAATGCTGAGCGCCTCTCTCAGC	1337
DB	1204 CTGAGCCCCCTACAGAAAGCTTAAGCATGGACCCCGGCAATGCTGAGCGCCTCTCTCAGC	1263
QY	1338 TTGGACGCTGTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAAG	1397
DB	1264 TTGGACGCTGTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAAG	1323
QY	1398 CGAGAGCGGATGGTGTCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT	1457
DB	1324 CGAGAGCGGATGGTGTCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT	1383
QY	1458 AAGACGAAGCAAAAGAACTGGAGGCAAGATGTTGGCCCGAGAGGCTGAGGAAAAGGAG	1517
DB	1384 AAGACGAAGCAAAAGAACTGGAGGCAAGATGTTGGCCCGAGAGGCTGAGGAAAAGGAG	1443
QY	1518 AACCATTTGCCCAANTGTGA 1538	
DB	1444 AACCATTTGCCCAANTGTGA 1464	
RESULT 9		
LOCUS	AR210052 1115 bp DNA linear PAT 20-JUN-2002	
DEFINITION	Sequence 1 from patent US 6387644.	
ACCESSION	AR210052	
VERSION	AR210052.1 GI:21512182	
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 1115)	
AUTHORS	Beraud,C.	
TITLE	Motor proteins and methods for their use	
JOURNAL	Patent: US 6387644-A 1 14-MAY-2002;	
FEATURES	Location/Qualifiers	
source	1..1115	
	/organism="unknown"	
	/mol_type="unassigned DNA"	
ORIGIN		
Query Match	72.3%; Score 1112.4; DB 6; Length 1115;	
Best Local Similarity	99.9%; Pred. No. 3.9e-279;	
Matches 1113; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 ATCAGCGCGCGCGGCTCGACGACAGAGCGGCGAGATGCGGCGAGCTTCAGCGG	60
DB	1 ATCAGCGCGCGCGGCTCGACGACAGAGCGGCGAGATGCGGCGAGCTTCAGCGG	60
QY	61 CGCGCATCTCAGAGCTGTGCTGTGCGGTAAAGCAAGATTGAGACTACTCGTCTCCAC	120
DB	61 CGCGCATCTCAGAGCTGTGCTGTGCGGTAAAGCAAGATTGAGACTACTCGTCTCCAC	120

DEFINITION	Mus musculus kinesin family member 22, mRNA (cdna clone MGC:6456 IMAGE:2615715), complete cds.
ACCESSION	BC003427
VERSION	BC003427.1 GI:13097359
KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2086)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F., Dickenson,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toehiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.D., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Viallon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Snailus,D.E., Schnarch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBLISHED	12477932
REFERENCE	2 (bases 1 to 2086)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-rc@mail.nih.gov">cgapbs-rc@mail.nih.gov</a> Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAC Plate: 5 Row: m Column: 7 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704181.
SOURCE	Location/Qualifiers 1. .2086 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:6456 IMAGE:2615715" /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal Carcinoma. 5 month old virgin mouse." /clone_lib="NCI CGAP_Mam6" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6" 1. .2086 /gene="Kif22"
gene	

BC003427 2086 bp mRNA linear ROD 29-JUN-2004

/note="synonym: Kid"  
 /db\_xref="LocusID:110033"  
 /db\_xref="MGI:109233"  
 20. .2002  
 /gene="Kif22"  
 /codon\_start=1  
 /product="kinesin family member 22"  
 /protein\_id="AAH03427.1"  
 /db\_xref="GI:13097360"  
 /db\_xref="LocusID:110033"  
 /db\_xref="MGI:109233"  
 /translation="MSLRKTKPORREMASATSPGRCSKGLGRPPILARVRVAVR  
 LRPMDGETAKELPCVRAIDSCGLEVANKKQETLKYQFDAPFYGSKSQOEVYVGS  
 VQPLRLHLLSQNASVLAYGPTGAGKTHWLGSPQGVIPRALMDLLQLARESAG  
 RPWDVSVAMYLEIYQKLDLDPASGLVIREDCRGNLIIFGLTKYLPITSFDEPQ  
 HFDPASNRVAVLQNRSSHAVLLVQDRRLTPFRQREGKLDLIDLAGSDEN  
 RRTNGOIRKESGAINISFLVGLKVVDALNOGLPRIPIYRDSKLRLDLOSLGSAHS  
 ILIANIAPERFYDDTISALNFTARSVINRPTNESLOPHALAPVKLSOKELLGPS  
 EAKKAPPEESTGPESTAPASOKLSLLOKLSNMDPAMLENLLSMERLLGSGS  
 QGTPLLNTPKREVRMLKTVEEKLEIERLKMKELEAKVLAQEAADPPREKENTPTI  
 LQPPASVSGTVKPLKAVNMPQRIQKQRESSNOIQLLKKGPKRLKLPESAEVAK  
 DEVDWEVQISPELLAHGRKKLLDLDLNEGSAEELRSLQRIQOKKAQLIVGWRELHGFPS  
 EVDLEQVEGISGRQVESFLKANLLSLAASQHSGPS"

ORIGIN

Query Match 69.9%; Score 1074.8; DB 10; Length 2086;  
 Best Local Similarity 83.6%; Pred. No. 2.7e-289; Mismatches 238; Indels 0; Gaps 0;  
 Matches 1214; Conservative 0;

Qy	59	GGCGGCGATCTCAGAGCTGTGCTCGCTGCTGAGCAAGATTGAGCTACTCGTCGTC	118
Db	61	GGCGTGGCCACCTCAGGACCTGCGCCGCTGTGTAAGCAAGAGGGCTTGGCCGGCTCC	120
Qy	119	ACCTCCAGCTCGCTAAGGGTGGCTGCGAGCTCGGCCATTTGTGATGCAACAGCGGG	178
Db	121	GCCTCTAGCTCGAGTAAGGGTGGCTGCGGCTGCGCCCGTTTATGATGAGAGACAGA	180
Qy	179	AGCAAGTGATCCCGCTGTGCGGGGATGGAGAGCTGCTCTAGAGATTGCTACTG	238
Db	181	AGCGAAGGAGTCCCGCTGTGCGAGCATAGACAGCTGCTCTTGAAGTGGCTAACTG	240
Qy	239	GAGGAACACACAGAGAGCTCTCAATACCAAGTTTGTATGATGCGCTTATGGGAGAGAGTAC	298
Db	241	GAAGAAATACAGAGAGCTCAATATCAGTTTGTATGCTTCTATGCGGAGAGACAC	300
Qy	299	TCAGCAGACATCTATGAGCTTCAGTGCGACCCATCTTAAGGCACTTGTGGAAGGCA	358
Db	301	TCAGCAGAACTCTATGAGCTTCAGTGCGACCCATCTTAAGGCACTTGTGGAAGGCA	360
Qy	359	GAATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGAGAGCGCACCAATGCTGGG	418
Db	361	GAATGCCAGTGTGCTTGCCTATGAGCCCTATGAGGCGAGGAGAGACACCAATGCTGGG	420
Qy	419	CAGCCAGAGCAACTCGGGGTGATCCCGGGCTCTCATGAGCTCTGAGCTCACAAG	478
Db	421	CAGCCAGAGCAACTCGAGTGTATCTCGGGCTCTCATGAGCTCTGAGCTCACAAG	480
Qy	479	GGAGAGGGTCCGAGGGCGGCGCATGGCGCTTCTGTGTCACCATGCTTACCTTAGAGAT	538
Db	481	GGAGAGAGTGTCTGAGGAGCGGCCATGGAGCTTCTGTGCTATGCTTATTTAGAGAT	540
Qy	539	CTACCCAGGAGAGTATAGACCTCTGAGCCCTGCTTGGGAGACCTGCTATTCGAGA	598
Db	541	CTACCCAGGAGAGTATAGACCTCTTGGATCTGTCATCAGGAGACCTGCTGATTCGCGA	600
Qy	599	AGACTCGGGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGC	658
Db	601	AGACTCGGGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTGC	660
Qy	659	TGATTTGAGCGGCACTTCTGCGCAGCGAGTGGAAATCGAGCTGTAGAGCCACCGGCT	718
Db	661	TGATTTGAGCGGCACTTCTTCCAGCGAGTAGAAATCGAGCTGTAGAGCCACCGGCT	720

Qy	719	CAACGAGCGCTCTCCGCGAGTCATGCTGTGCTCTGCTCAAGTGACGAGCGGGAACG	778
Db	721	TAACGAGCGCTCTCCGCGAGTCATGCTGTGCTCAAGTGAGTAGTCAAGCTGAACG	780
Qy	779	TTTGGCCCGCATTTTCGCGAGCGAGGAGGAAAACCTCTACCTGATTGACTTGGCTGGT	838
Db	781	TTTGGCTCCATTTTCGCGAGCGGGAAGGAAAACCTCTACCTGATTGATTGGCTGGT	840
Qy	839	GGACCAACCGCGGACAGGCAACAGGGCCTTCGGCTTAAAGAGAGTGAGGAGCATCAAC	898
Db	841	GGACCAACCGCTCGACAGGCAACAGGGCATTTCGGCTCAAAAGAGAGCGGAGCATCAAC	900
Qy	899	CTCCCTGTTTGTCTCTGGGCAAGTGTAGATGCGCTGAATCAGGGCCTCCTCGTGACC	958
Db	901	CTCCCTGTTTGTCTCTGGGCAAGTGTGTGATGATTAAACAGGGCCTCCTCGCATACC	960
Qy	959	TTATCGGAGACAGCAAGCTCACTCGCCCTATTTCAGGAGCTCTCTGGGTGGCTCAGCC	1018
Db	961	ATACCGGAGACAGCAAGCTCACTCGCTGCTGCGAGGACTCTCTGGAGGCTCAGCTCAT	1020
Qy	1019	TATCCTTATTGGCAACATTGCCCTCTGAGAGACGTTCTACTAGACACAGTCTCCGCACT	1078
Db	1021	CATCCTCATTTGCCAACATTGCTCTGAAAGACGTTTTTACAGGATACAAATTTGAGATT	1080
Qy	1079	CAACTTGTGCTGCCAGGTCCAAAGAGGTGATCAATCGGCCCTTTTACCAATCAGAGCTGCA	1138
Db	1081	GAATCTCAGTGTGAGTCCAAAGAGGTGATTAACCGGCCCTTTTACCAATGAGAGTCTACA	1140
Qy	1139	GCCTCATGCTTGGGACCTGTTAAGCTGCTCAGAAAAGAAATTTGTTGGTCCACAGAGGC	1198
Db	1141	GCCTCATGCTTGGCACCCTGTTAAGCTGCTCAGAAAAGAAATTTGTTGGTCCAGAGGC	1200
Qy	1199	AAAGAGAGCGCGGCGCTCAGGAAAGAGAGATTTGAGAGCCCTTGGAGCCATTTGAGCTCC	1258
Db	1201	AAAGAGAGCGGCGCTCAGGAAAGAAATCAACTGGGAGCCCTTGGAGCTTACAGAGCTCC	1260
Qy	1259	AGCTCTGCTCCAGAACTCAGCCCTTACAGAACTTAAAGAGCTTAAAGCAATTTGATCCAGGCAT	1318
Db	1261	AGCTCTGCTCCAGAACTTAAAGCTTAAAGCAATTTGATCCAGGCAT	1320
Qy	1319	GCTGGAGCGCTCTCTCAGCTTGGACCGTCTGCTTGGCTTCCAGGGGAGCCAGGGGGCCCC	1378
Db	1321	GTTGGAGAACCTCTCTGATGATGAACGTTTGTGCTGGTTCCTCAGGAGAACCAAGGACCCC	1380
Qy	1379	TCGTTGATGATACCCAAAGAGAGCGGATGCTTAAAGAAAGAGTAAAGAGAGAGGA	1438
Db	1381	CCTACTGAATACCCAAAGAGAGCGAATGCTTAAAGAAAGAGTAAAGAGAGAGGA	1440
Qy	1439	CCTAGAGATTGAGAGGCTTAAAGAGAGCAAAAGAACTGGAGGCCAAGATTTGGGCCA	1498
Db	1441	CTTAGAGATTGAGAGGCTTAAAGATGAAGCAAAAGAACTGGAGGCCAAGATTTGGGCCA	1500
Qy	1499	GAAGGCTGAGGA 1510	
Db	1501	GGAGGACCGGA 1512	

RESULT 11  
 AR210054  
 LOCUS  
 DEFINITION Sequence 5 from patent US 6387644.  
 ACCESSION AR210054  
 VERSION AR210054.1 GI:21512185  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1041)  
 AUTHORS Beraud, C.  
 TITLE Motor proteins and methods for their use  
 JOURNAL Patent: US 6387644-A 5 14-MAY-2002;  
 FEATURES Location/Qualifiers  
 source 1. .1041

ORIGIN  
Query Match 67.3%; Score 1035.4; DB 6; Length 1041;  
Best Local Similarity 99.9%; Pred. No. 4.9e-259;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	78	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCCAGCTCGCTAAGG	137
DB	4	GGTCGCTGTCGGCTAAGCAAGATTGGAGCTACTCGTCTGCTCCACCTCCAGCTCGCTAAGG	63
QY	138	GTGGCTGTGCGACTGCGGCCATTGTGTGATGAACACAGCGGGAGCAAGTATCCCCCTGT	197
DB	64	GTGGCTGTGCGACTGCGGCCATTGTGTGATGAACACAGCGGGAGCAAGTATCCCCCTGT	123
QY	198	GTGCGGGGATGAGAGCTGCTCTTAGAGATTGCTTAAGTGAAGAACACAGGAGACT	257
DB	124	GTGCGGGGATGAGAGCTGCTCTTAGAGATTGCTTAAGTGAAGAACACAGGAGACT	183
QY	258	CTCAATACCACTTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	317
DB	184	CTCAATACCACTTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243
QY	318	GTTTCAGTCAGCCCATCTAAGGCACTTGTCTGGAAGGCGAGATGCCAGTGTGCTGCC	377
DB	244	GTTTCAGTCAGCCCATCTAAGGCACTTGTCTGGAAGGCGAGATGCCAGTGTGCTGCC	303
QY	378	TATGACCCACAGGAGCTGGGAAGCGCACAAATGCTGGGAGCGCCAGAGCAACTGGG	437
DB	304	TATGACCCACAGGAGCTGGGAAGCGCACAAATGCTGGGAGCGCCAGAGCAACTGGG	363
QY	438	GTGATCCCGCGGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGGTCCCGAGGCG	497
DB	364	GTGATCCCGCGGGCTCTCATGACCTCTGAGCTCACAAGGAGGAGGGTCCCGAGGCG	423
QY	498	CGCCATAGGCGCCCTTCTGTCTACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA	557
DB	424	CGCCATAGGCGCCCTTCTGTCTACCATGCTTACCTAGAGATCTACAGGAGAGGTATTA	483
QY	558	GACTCTCGGACCTGCTTCGGGACCTGTTAATCCGAGAGACTGCGGGGGAATATC	617
DB	484	GACTCTCGGACCTGCTTCGGGACCTGTTAATCCGAGAGACTGCGGGGGAATATC	543
QY	618	CTGATTCGGGGTCTCTCCAGAAAGCCATCAGTAGCTTGTGATTTTGTAGCGGCACTTC	677
DB	544	CTGATTCGGGGTCTCTCCAGAAAGCCATCAGTAGCTTGTGATTTTGTAGCGGCACTTC	603
QY	678	CTGCCAGCGAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC	737
DB	604	CTGCCAGCGAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCTCCGC	663
QY	738	AGTCATGCTGCTCTGCTGAGTGGACGACGCGGAGCGTTTGGCCCATTTGCCAG	797
DB	664	AGTCATGCTGCTCTGCTGAGTGGACGACGCGGAGCGTTTGGCCCATTTGCCAG	723
QY	798	CGAGAGGGAATCTTACTGATTCACCTTGGCTGCTGAGGCAACCGCGGACAGGCG	857
DB	724	CGAGAGGGAATCTTACTGATTCACCTTGGCTGCTGAGGCAACCGCGGACAGGCG	783
QY	858	AACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	917
DB	784	AACAGGGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGC	843
QY	918	AAAGTGTAGTGCCTGAATCAGGGCTCTCCCTGCTGATCTTATCGGGACGGAAGCTC	977
DB	844	AAAGTGTAGTGCCTGAATCAGGGCTCTCCCTGCTGATCTTATCGGGACGGAAGCTC	903
QY	978	ACTGCGCTATTGACGACTCTCTGGGTGGCTCAGCCGACAGTATCTTATGCGCAACT	1037
DB	904	ACTGCGCTATTGACGACTCTCTGGGTGGCTCAGCCGACAGTATCTTATGCGCAACT	963
QY	1038	GCCCTGTAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCC	1097

Db	964	GGCCCTGTAGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCC	1023
Qy	1098	AAGGAGGTGATCAATCG	1114
Db	1024	AAGGAGGTGATCAATTTG	1040
RESULT 12	AC101919/c	151041 bp	DNA linear HTG 14-MAR-2003
LOCUS	Mus musculus clone RP24-158A6, WORKING DRAFT SEQUENCE, 4 unordered pieces.		
ACCESSION	AC101919		
VERSION	AC101919.4	GI:28951338	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 151041)		
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.		
TITLE	Mus musculus, clone RP24-158A6		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 151041)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 151041)		
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		



Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 14, 2003 this sequence version replaced gi:28631351.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information

Center project name: L17745

Center clone name: 158 A.6

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 150558 bases at least Q40

Consensus quality: 150654 bases at least Q30

Consensus quality: 150687 bases at least Q20

Insert size: 151000; agarose-fp

Insert size: 150741; sum-of-contigs

Quality coverage: 11.8 in Q20 bases; agarose-fp

Quality coverage: 11.8 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 33517: contig of 33517 bp in length  
 33518 33617: gap of 100 bp  
 33618 42736: contig of 9119 bp in length  
 42737 42836: gap of 100 bp  
 42837 84394: contig of 41558 bp in length  
 84395 84494: gap of 100 bp  
 84495 151041: contig of 66547 bp in length.

## FEATURES

## source

1. 151041  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP24-158A6"  
 /clone\_lib="RPCI-24 Male Mouse BAC"

## misc\_feature

1. 33517  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left

## misc\_feature

33618..42736  
 /note="assembly\_fragment"

## misc\_feature

42837..84394  
 /note="assembly\_fragment"

## misc\_feature

84495..151041  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right

## ORIGIN

Query Match 57.68; Score 885.2; DB 2; Length 151041;  
 Best Local Similarity 78.2; Pred. No. 1.1e-219;  
 Matches 1136; Conservative 0; Mismatches 288; Indels 28; Gaps 5;  
 57 GCGCGCGCGATCTCAGGAGTGTGCTGTCGGCTAAGCAAGATGGAGCTACTCGTCGT 116  
 65725 GCGTCACCCACCTCAGACCGGGTCTTGT---ATAAGCAAGGTAGGATGGCTGTCTG 65669  
 117 CCACCTCAGCTCGCTAAGGTTGGCTGTGGACTCGGCCATTGTGGATGGAACAGCG 176  
 65668 CTGCTTCTAGCTGGATGAAGGGTGGCTATATATATGTCCTGTTTGTGGATGAAGACA 65609

QY 177 GGAGCAAGTGATCCCCCTGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAAC 236  
 DB |||||  
 65608 GAAGCGAAGGAGCACCCTGTGTACAGGCCATAGACGCTGCTCTCTTGAAGTGGCTAAC 65549  
 QY |||||  
 237 TGGAGGAACACACAGGAGACTCTCAAATACCACTTTGATGCCCTTCTATGGGAGAGAGT 296  
 DB |||||  
 65548 TGGAGGAATACCAGGAGACTATCAAATATCACTTTGATGCCCTTCCATGGAGAGAGC 65489  
 QY |||||  
 297 ACTCAGCAGGACATCTATGACAGGTTAGTGACGCCCATCTAAGGCATCTTCTGGAAGGG 356  
 DB |||||  
 65488 ACTCAGTGGGACATCTATGTAGGTTTAGTACAGCCCATATAAGGAACCTTCTGGAAGGG 65429  
 QY |||||  
 357 CAGAATGCCAGTGTCTTGCTATGACCCACAGGAGCTGGGAAGCGCACACAATCCTG 416  
 DB |||||  
 65428 CAGAATGCCAGTGTCTTGCTATGACCCACTGGGGCAGGGAAGA--CACACAATGTTG 65371  
 QY |||||  
 417 GGCAGCCCAAGAGCAACCTGGGGTGATCCCGCGGGCTCTCATGGACCTCTCTGCACTCACA 476  
 DB |||||  
 65370 GGCAGCCCAAGAGCAACCTGGGGTGATCCCTCAGACTCTCCAGACCTCTCTGCACTTAA 65311  
 QY |||||  
 477 AGGAGAGGGTCCGAGGCGCGGCATGGGCCCTTCTGTCAACATGCTCTTACCTAGAG 536  
 DB |||||  
 65310 AGGGAAGAGAACTCTGAGGGCTGGCCATGGGACATTTCTGTGCTTGTCTATTTAGAT 65251  
 QY |||||  
 537 ATCTACCAGGAGAGGTATTAGACCTCTCGACCTCTTCGGGAGACCTGGTAATCCGA 596  
 DB |||||  
 65250 CTCTAACAGGAATAGTATTAGACCTCTTGGATCTCGATCAGAAGACCTAGTGTCCGA 65191  
 QY |||||  
 597 GAAGACTGCCGGGGGAATATCTTGATTCGGGGTCTCTCCAGGAAGCCCATCAGTAGCTTT 656  
 DB |||||  
 65190 GAAGGCTGACGAGGGAATCTCTGATTCAGAGCTCTCACAGAAAGCCCATCCTAGTTT 65131  
 QY |||||  
 657 GCTGATTTGAGCGGCATCTCTGCCAGCCAGTCGAAATCGGACTGTAGAGGCCACCCGG 716  
 DB |||||  
 65130 TCTGACTTCGAGAAGCACTCTCTTAGTCAGTAGAATCGAGTTGTAGGAGCTACTCTG 65071  
 QY |||||  
 717 CTCAACACAGCGCTCTCCGCGCAGTCATGCTGCTCTGCTCAAGGTGGACGAGCGGAA 776  
 DB |||||  
 65070 CTTTAAACAGCGCTCTTGGCGGTAGTCATGACGCGCTCTTGGTCAAGGTAGATCAGCATGAA 65011  
 QY |||||  
 777 CGTTTGGCCCCATTTCCGCCAGCGAGAGGGAAAACTCTACTGATTTGACTTGGCTGGGTCA 836  
 DB |||||  
 65010 CGTTTGACTCCAATTCGCCAGTGGGAAGGAAAACTCTACTTATTGATTTGGCTGGTTCA 64951  
 QY |||||  
 837 GAGCAACACC-----GGCGCACAGGCAACAGGGCCCTTCGGCTAAAGAGAGTGGAGCAT 892  
 DB |||||  
 64950 GAGGCAACCGTCAGTCGCAAGGCAACCAAGGCAATTTCTGCTCAAAAGAGAGTGGAGCAT 64891  
 QY |||||  
 893 CAACACCTCCCTGTTGTCTGGGCAAGTGGTAGATGCGCTGAATCAGGGCCCTCCCTCG 952  
 DB |||||  
 64890 CAACACCTCCCTGTTGTACTGGGCAAGTAGTGGATGATTAACCAAGGCCCTCCCTCA 64831  
 QY |||||  
 953 TGTACCTTATCGGACAGCAAGCTCACTCGCCCTATTTCAGGACTCTCTGGGTGGCTCAGC 1012  
 DB |||||  
 64830 TATACCATACAGGACAGCAGGCTCACTGCTGCTGTTGCAAGACTCTCTGAGAGGCTCAGC 64771  
 QY |||||  
 1013 CCACAGTATCTTATTGGCCAACTTCCCGCTGAGAGACGCTTCTACCTAGACACAGTCTC 1072  
 DB |||||  
 64770 TCATAGCATCTCTCATTTGTGAGCATTTGCTCTCTGAAAGATGTTTTTACCTGGATCAATCTC 64711  
 QY |||||  
 1073 CGCACTCAACTTTGCTGCCAGGTCCCAAGGAGGTGATCAATCGGCCCTTTTACCAATGAGAG 1132  
 DB |||||  
 64710 AGCATTTAACTTCACTGCTAGGTCCCAAGGAGGTGATTAACAGGCTTTTACCAAGGAG 64651  
 QY |||||  
 1133 CTTGACGCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGCTGGTCCACC 1192  
 DB |||||  
 64650 TGTACAACTCTCAGGCTTGACACCTGTTAAGCTGTCTCAGAAAGATCTGCTAGGTTTATC 64591  
 QY |||||  
 1193 AGAGGCAAGAGAGCCGAGGCGCTCAGGAAGAGGAGATGGGAGCCCTCAGGCCCTAGC 1252  
 DB |||||  
 64590 AGAGGCAAGAAAGCCAAAGGCGCTCAAGAAAGAAATCAACTGGGAGCCCTGAGTCTATAGC 64531  
 QY |||||  
 1253 AGCTCCAGCCTCTGCTCTCCAGAAAATCTCAGCCCCCTACAGAAGCTAAGCAGCATGGACCC 1312



```

Db      64530 ATCTCCAGCATCTGCTTCCAGAAA-----CTAAGCAATATGATCT 64489
QY      1313 GCCCATGCTGGAGCGCCTCTCAGCTTGACCGTCTGCTTGCTCCAGGGAGCCAGG 1372
Db      64488 AGCCATGTTGGGAAGCTCTCAGTATGGAATGTTTGTGCTTCAAGGGAGCCAGAG 64429
QY      1373 GCCCCTCTGTTGAGTACCCAAAGCGAGAGCGGATGGTGTAAATGAAGACAGTAGTAAGA 1432
Db      64428 GAACCTCTACTGAATACCCCAA--CGAGAGCGAATGGTGTCTATTAGACGGTGGAGGA 64370
QY      1433 GAAGACCTTAGATGTAGAGCTTAAAGCAGAAAGCAAAAGAACTGGAGGCCAAGATGTT 1492
Db      64369 CAAGATATTGGAGATTGAGAGCTTAAAGATGAAGCAAAAGAACTGGAAGCCCGCTGCT 64310
QY      1493 GCCCCAGAGGC 1504
Db      64309 GTCCAGGAAGC 64298

```

```

RESULT 13
AC101752/c
LOCUS   Mus musculus chromosome 1, clone RP24-344C18, complete sequence.
DEFINITION
ACCESSION AC101752
VERSION   AC101752.14 GI:46518680
KEYWORDS HTG.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 164759)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 1, clone RP24-344C18
Unpublished
2 (bases 1 to 164759)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
ChoepeI,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cook,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lanazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164759)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., ChoepeI,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,

```

```

Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164759)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
ChoepeI,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-APR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L17344
Center clone name: 344_C_18
-----
Location/Qualifiers
1. 164759
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clones="RP24-344C18"
/clone_lib="RP24-344C18"
1. 1655
/repeat_region
misc_feature
1. 4
/feature="clone boundary"
clone_end:SP6
site:MbolI
1767..1787
/rpt_family="AT_rich"
2115..2157
/rpt_family="AT_rich"
2236..2258
/rpt_family="AT_rich"
3252..3420
/rpt_family="Lx7"

```

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## FEATURES

## source

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

```
repeat_region 3464. .3588  
/rpt_family="Lx7"  
repeat_region 3605. .3666  
/rpt_family="CT-rich"  
repeat_region 3673. .3731  
/rpt_family="TTC)n"  
repeat_region 3732. .3911  
/rpt_family="TTC)n"  
repeat_region 4000. .4180  
/rpt_family="CT-rich"  
repeat_region 4202. .4363  
/rpt_family="Lx7"  
repeat_region 4529. .4765  
/rpt_family="MTE"  
repeat_region 7234. .7442  
/rpt_family="B3"  
repeat_region 8445. .8509  
/rpt_family="T-rich"  
repeat_region complement(8512. .8591)  
/rpt_family="MLT11"  
repeat_region 9336. .9632  
/rpt_family="MLT1J"  
repeat_region complement(9800. .11396)  
/rpt_family="L1_MM"  
repeat_region complement(11397. .17527)  
/rpt_family="L1P"  
repeat_region complement(18235. .18777)  
/rpt_family="Lx"  
repeat_region .18775. .18890  
/rpt_family="Lx"  
repeat_region 19260. .19949  
/rpt_family="RMER19A"  
repeat_region 20118. .20297  
/rpt_family="MLT1D"  
repeat_region complement(20386. .20639)  
/rpt_family="Lx2"  
repeat_region 20634. .20839  
/rpt_family="Lx4"  
repeat_region 20876. .21793  
/rpt_family="RLTFR13D"  
repeat_region 21799. .21907  
/rpt_family="RMER13A"  
repeat_region 21918. .22433  
/rpt_family="Lx4"  
repeat_region 22591. .22651  
/rpt_family="(TG)n"  
repeat_region 22733. .24486  
/rpt_family="L1_MM"  
repeat_region 24494. .24850  
/rpt_family="L1_MM"  
repeat_region 27139. .27166  
/rpt_family="AT-rich"  
repeat_region 27737. .27834  
/rpt_family="MLT1H"  
repeat_region 28328. .28506  
/rpt_family="MLT1K"  
repeat_region 28513. .28541  
/rpt_family="AT-rich"  
repeat_region 28813. .28858  
/rpt_family="AT-rich"  
repeat_region 29627. .30186  
/rpt_family="L1_MM"  
repeat_region 30465. .30578  
/rpt_family="T-rich"  
repeat_region complement(30581. .34457)  
/rpt_family="L1_MM"  
repeat_region complement(34455. .35895)  
/rpt_family="L1"  
repeat_region complement(36134. .36736)  
/rpt_family="Lx4"  
repeat_region 36749. .36920  
/rpt_family="Lx2"  
repeat_region 37617. .37751
```

```
repeat_region /rpt_family="(TA)n"  
37958. .37986  
/rpt_family="(GAA)n"  
Query Match 57.6%; Score 885.2; DB 10; Length 164759;  
Best Local Similarity 78.2%; Pred. No. 1.1e-219;  
Matches 1136; Conservative 0; Mismatches 288; Indels 28; Gaps 5;  
QY 57 GCGCGCGGATCTCAGAGCTGCTGCTGCGGCTAAGCAAGATTGGAGCTACTCGTCGT 116  
Db 146137 GCCTCAGCCACCTCAGGACCGGCTTTTGT- --ATAAGCAAGGTAGGATTTGGCTGCT 146081  
QY 117 CCACCTCCAGCTCGCGTAAGSGTGGCTGCGAGCTGGGCCATTTTGGATGGAACAGCG 176  
Db 146080 CTGCTTCTAGCTGGAGTAAGGGTGGCTATATAATTGTGCTTGTGTTGTAAGAGACACA 146021  
QY 177 GGAGCAAGTGATCCCTCTGTGCGGGGCGATGACAGCTGCTCTCTAGAGATTGCTTAAC 236  
Db 146020 GAAGCGAAGGAGCACCCCTGTGTACGAGCCATAGACAGCTGCTCTCTTGAAGTGGCTAAC 145961  
QY 237 TGGAGGAACCAAGAGAGACTCTCAAAATACAGCTTTGATGCTCTTATGGGGAGAGAGT 296  
Db 145960 TGGAGGAATAATACAGGAGACTATCAAAATATCAGTTTGTATGCTCTTCCATGGAGAGAGC 145901  
QY 297 ACTCAGCAGGACATCTATGAGGTTTCAGTGACGCCCATCTAAGGCACCTTGTGGAGGG 356  
Db 145900 ACTCAGTGGGACATCTATGAGGTTTATGACGCCCATCTAAGGAACCTTGTGGAGGG 145841  
QY 357 CAGAATGCCAGTGTCTTGCTATGGACCCACAGGAGCTGGGAGAGCGACAACAATGCTG 416  
Db 145840 CAGAATGCCAGTGTCTTGCTATGGACCGACTGGGCGGAGGA- -CACACAATGTTG 145783  
QY 417 GGCAGGCCAGAGAACTCTGGGTGATCCCGGGGCTCTCATGACCTCTGAGCTCACA 476  
Db 145782 GGCAGGCCAGAACAGGCTTGGAGTGATTCCTCAGACTCTCCAGACCTCTGCACTAAC 145723  
QY 477 AGGAGGAGGGTCCGAGGCGCGCATGGGCCCTTTCTGTGTCACCATGCTCTTACCTAGAG 536  
Db 145722 AGGAGAGAGATGCTGAGGCTGGCCATGGGACATTTCTGTGCTTGTGCTATTATAGAT 145663  
QY 537 ATCTACAGGAGAAAGTATTAGACCTCTTGGACCTCTGCGGAGACCTTGGTAATCCGA 596  
Db 145662 CTCTAACAGGAATAGTATTAGACCTCTTGGATCTCTGCACTCAGAGACCTAGTGTTCGA 145603  
QY 597 GAGAGTCCGCGGGGAATATCTGATTCGGGTCTCTCCAGAGCGCATCAGTAGCTTT 656  
Db 145602 GAAGGCTGACGAGGGAATCTCTGATTCAGGCTCTCACAGAGGCCCATCTACTTTTC 145543  
QY 657 GCTGATTTTGAGCGGCACCTTCTGCGCAGCGAGTGGAAATCGGACTGTAGGAGCACCGG 716  
Db 145542 TCTGACTTCGAGNAGCACTTCTCTTAGTCAGTAGAATCGAGTGTAGGAGTACTCTG 145483  
QY 717 CTCAACAGCGCTCTCTCCGCGAGTCACTGTGCTCTCTGCTCAAGGTGGACCGCGGAA 776  
Db 145482 CTTAACAGCGCTCTTGGCGTAGTCTATGAGCGCTCTTGGTCAAGGTAGATCAGCATGAA 145423  
QY 777 CGTTTGGCCCCATTTGCCAGCGAGGGAATCTACTGATGACTTGGCTGGGTGCTCA 836  
Db 145422 CGTTTGACTCCATTTGCCAGTGGGAAGAAATCTACTGATGATTTGGCTGGGTGCTCA 145363  
QY 837 GAGGACAAAC- --GGCGCACAGGCAACAAAGGGCTTCGGCTAAAAGAGAGTGGAGCCAT 892  
Db 145362 GAGGACAAACCGTCAGTCGACAGGCAACCCAGGGCATTTCTGCTCAAGAGAGTGGAGCCAT 145303  
QY 893 CAACACTCTCTGTTTGTCTCTGGGCAAAAGTGGTAGATGCGCTGAATCAGGCGCTCCCTCG 952  
Db 145302 CAACACTCTCTTGTGTACTGGGCAAAAGTGGTAGATGCAATTAACCAAGGCTCCCTCA 145243  
QY 953 TGTACCTTATCGGACAGCAAGCTCACTCGCTATTGCAAGGACTCTCTGGTGGCTCAGC 1012  
Db 145242 TATACCATACAGGACAGGAGGCTCACTCGTCTGTTGCAAGACTCTCTGAGAGGCTCAGC 145183  
QY 1013 CCACAGTATCTTTATTGCCAAACATTTGCCCTTGAGAGACGCTTCTTACCTAGACACAGTCTC 1072
```

```
Db 145182 TCATAGATCCTCATTTGTCAGCATTCCTCTGAAGATGTTTTTACCTGGATCAATCTC 145123
QY 1073 CGCACTCAACTTTGCTGCGCAGGTCGAAGAGGTGATCAATCGGCCCTTTTACCAATCAGAG 1132
Db 145122 AGCATTAACCTTCACTGCTAGGTCGAAGAGGTGATTAACCAAGCTTTCCAAACGAG 145063
QY 1133 CTGCGAGCTCATGCTCTGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACC 1192
Db 145062 TGTACAACTTCAGGCTTGACACCTGTTAAGCTGTCTCAGAAAGATCTGCTAGGTTTATC 145003
QY 1193 AGAGGCAAGAGAGCGCGGCGCTCAGGAAGAGAGATTTGGAGCCCTTGGAGCCCATGGC 1252
Db 145002 AGAGGCAAGAGAGCGCGGCGCTCAGGAAGAGATCAACTGGAGCGCTGAGTCTATAGC 144943
QY 1253 AGCTCAGAGCTCTGCTCTCCAGAACTCACTGCCCTTACAGAGCTTAAGCAGCATGGACCC 1312
Db 144942 ATCTCAGCATCTGCTTCCAGAA-----CTAAGCAATATGATCT 144901
QY 1313 GGCCATGCTGGAGCGCTCTCTCAGCTTGGACCGTCTGCTTCCCTCCAGGGAGCCAGGG 1372
Db 144900 AGCCATGTTGGGAAAGCTCTCTGATGTAATGTTTCTGGTGTCTCAAGGGAGCCAGAG 144841
QY 1373 GGCCCTCTGTTGAGTACCCCAAGCAGAGCGGATGCTGCTTAATGAAGACAGTAGAAGA 1432
Db 144840 GAACCTCTACTGAATACCCCAA--CGAGAGCGAATGGTGTCTCATTAAGACGCTGGAGGA 144782
QY 1433 GAAGGACCTTAGAGATGAGAGCGCTTAAGACGAGCAAGAAAGAACTGAGGCCCAAGATGTT 1492
Db 144781 CAGATATTGGAGATTGAGAGGCTTAAGATGAAGCAAGAAAGAACTGGAAGCCAGGTGCT 144722
QY 1493 GGCCCAAGAGGC 1504
Db 144721 GGTCCAGGAGC 144710
```

```
RESULT 14
AC101908/c
LOCUS
DEFINITION Mus musculus chromosome 1 clone RP24-275J1 map 1, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
AC101908
VERSION AC101908.4 GI:44199137
KEYWORDS HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEPIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 1, clone RP24-275J1
Unpublished
2 (bases 1 to 297639)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cook,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Hulme,W., Iliev,I.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Milnga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosett,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 297639)  
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
Anderson,S., Anderson,S., Atachchi,H.M., Barna,N., Bastien,V.,  
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,  
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Cotum,B.,  
DeArelano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,  
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Milnga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (18-JUN-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 27, 2004 this sequence version replaced gi:31880232.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@broad.mit.edu  
----- Project Information  
Center project name: Li7732  
Center clone name: 275\_J\_1  
-----

TITLE  
JOURNAL  
COMMENT

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 63747: contig of 63747 bp in length  
\* 63748 63847: gap of 100 bp  
\* 63848 73892: contig of 10045 bp in length  
\* 73892 73992: gap of 100 bp  
\* 73992 154450: contig of 80458 bp in length  
\* 154450 154550: gap of 100 bp  
\* 154550 167808: contig of 13258 bp in length  
\* 167808 167909: gap of 100 bp  
\* 167909 176711: contig of 8803 bp in length  
\* 176711 176811: gap of 100 bp  
\* 176812 233500: contig of 56689 bp in length  
\* 233500 233600: gap of 100 bp  
\* 233600 241222: contig of 7622 bp in length  
\* 241222 241322: gap of 100 bp  
\* 241322 297639: contig of 56317 bp in length.  
\* 241323 Location/Qualifiers  
1.297639  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"

```
/clone="RP24-275J1"
/clone_lib="RP24 Male Mouse BAC"

Query Match      48.4%; Score 744; DB 2; Length 297639;
Best Local Similarity 78.0%; Pred. No. 8.4e-183;
Matches 936; Conservative 0; Mismatches 255; Indels 9; Gaps 3;

Qy 57 GCGCGCGCGATCTCAGAGCTGTCGCTGTCGCGCTAAGCAAGATTTGAGCTACTCGTCGT 116
Db 1205 GCGTCAGCCACTCAGACCGGGTCTTTGT---ATAAGCAAGGTAGGATTTGGCTGTCGT 1149
Qy 117 CCACCTCCAGCTCGCTAAGGTGGCTGTCGACTCGGCCATTTTGGATGGAACAGCG 176
Db 1148 CTGCTTCTAGCTGAGTAAGGTGGCTATATAATTGTCCTGTTTGTGGATGAAGACA 1089
Qy 177 GGACCAAGTATCCCCCTGTGCGGGGATGACAGCTGCTCTAGAGATTGCTTAAC 236
Db 1088 GAAGCGAAGGACACCTCTGTACGAGCCATGACAGCTGCTCTTGAAGTGGCTAAC 1029
Qy 237 TGGAGGAACCAACAGAGACTCTCAATACAGTTTGTATGCTCTTATGGGAGAGAGT 296
Db 1028 TGGAGGAATACCAAGAGACTATCAATATCAGTTTGTATGCTCTTCCATGGAGAGAGC 969
Qy 297 ACTCAGCAGACATCTATGAGTTTCAAGTTCAGTGCAGCCCATCTAAGCACTTGTGGAGGG 356
Db 968 ACTCAGTGGGACATCTATGAGTTTGTATGAGTTCAGTGCAGCCCATCTAAGAACTTGTGGAGGG 909
Qy 357 CAGAAATGCCAGTGTCTTATGAGCCCAACAGAGCTGGGAGAGAGCAACAAATGCTG 416
Db 908 CAGAAATGCCAGTGTCTTATGAGCCCAACAGAGCTGGGAGAGAGCAACAAATGCTG 851
Qy 417 GGCAGCCAGAGCAACCTGGGCTGATCCCGGGGCTCTATGAGACTCTCTGAGCTCACA 476
Db 850 GGCAGCCAGAGCAACCTGGGCTGATCTCTCAGACTCTCCAGACCTCTCGACACTAACA 791
Qy 477 AGGAGAGGGTGGCGAGGGCGGCTTGTGAGGCTGGCCATTTCTGTCCACATGCTTACCTAGAG 536
Db 790 AGGAGAGAGATGCTGAGGGCTGGCCATTTGTGAGGCTGGCCATTTCTGTGCTATTTAGAT 731
Qy 537 ATCTACAGGAGAGAGTATTAGACTCTGTCGACCTGCTCGGAGAGCTGGTAATCCGA 596
Db 730 CTCTAACAGGAAATAGTATTAGACTCTTGGATCTTCTATAGTCACTAAATGCTGA 671
Qy 597 GAAGACTGCGGGGGAATATCTGATTCGCGGTCTCTCCAGAGAGCCATCAGTAGCTTT 656
Db 670 CGGCGAGAGAGGAGGAAATCTGATTCAGGCTCTCACAGAGAGCCATCAGTAGTTTC 611
Qy 657 GCTGATTTGAGCGGCACTCTCCGAGTCAATGCTGCTCTGCTCAAGGTGGACCGCGGAA 716
Db 610 TCTGACTTCGAGAGAGCACTTCTCTTCTAGTCAATGAGTAATGAGTTGAGGAGCTACTCTG 551
Qy 717 CTCAACAGAGGCTCTCTCCGAGTCAATGCTGCTCTGCTCAAGGTGGACCGCGGAA 776
Db 550 CTCAACAGAGGCTCTCTCCGAGTCAATGCTGCTCTGCTCAAGGTGGACCGCGGAA 491
Qy 777 CGTTTGGCCCATTTCCGAGAGAGGAGAAATCTTACTGATGATCTTGGCTGGGTGCA 836
Db 490 CGTTTGGATCTCCATTTCCGAGTGGGAGGAAATCTTACTGATGATTTGGCTGGGTGCA 431
Qy 837 GAGCAACACCC---GGCGCAGGCAACAGGGGCTTCCGCTTAAAGAGAGTGGAGCCAT 892
Db 430 GAGCAACACCCGTCAGTCGCAAGGCAACAGGGGCTTCTGCTTCAAGAGAGTGGAGCCAT 371
Qy 893 CAACACCTCTCTGTTTGTCTGGGCAAGTGGTATGCTGCTGATCAGGCGCTCTCCCTCG 952
Db 370 CAACACCTCTCTGTTTGTCTGGGCAAGTGGTATGCTGCTGATCAGGCGCTCTCCCTCG 311
Qy 953 TGTAACCTTATCGGAGACAGCAAGCTCCTGCGCTATTTGAGGAGCTCTCTGGTGGCTCAGC 1012
Db 310 TATACCAATACAGGAGACAGGAGCTCCTGCTGCTGTTGCAAGACTCTCTGAGGAGCTCAGC 251
Qy 1013 CCACAGTATCTTTATGCAACATTTGCGCCCTGAGAGGCTTCTTACCTAGACACAGTCTC 1072
```

```
Db 250 TCATAGCATCTCATTTGTCAGCATTTGCTCTCTGAAAGATGTTTTTCTCTGATACATCTC 191
Qy 1073 CGCACTCAACTTTGCTGCCAGGTCCCAAGGAGGTGATCAATCGCCCTTTTACCAATGAGAG 1132
Db 190 AGCATTTAAACTTCACTGCTAGGTCCCAAGGAGGTGATTAAACAGGCTTTTACCAAGGAG 131
Qy 1133 CTGCGAGCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGCTTGGTCCACC 1192
Db 130 TGTACAACTCTCAGGCTTGCACCTGTTAAGCTGTCTCAGAAAGATGCTGCTAGGTTTATC 71
Qy 1193 AGAGGCAAGAGAGAGCCGAGGCTCTGAGGAAGAGAGATGCGGAGCCCTGAGGCCATGGC 1252
Db 70 AGAGGCAAGAGAGAGCCGAGGCTCTGAGGAAGATCACTGGGAGCCCTGAGTCTATAGC 11

RESULT 15
AC102127 196674 bp DNA linear ROD 01-MAY-2004
LOCUS Mus musculus chromosome 1, clone RP23-202A19, complete sequence.
DEFINITION AC102127
VERSION AC102127.11 GI:46931440
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren,B., Nusbaum,C. and Lander,E.
1 (bases 1 to 196674)
Mus musculus chromosome 1, clone RP23-202A19
Unpublished
2 (bases 1 to 196674)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Choepl,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
McCarthy,M., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 196674)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boukhgalter,B., Camarata,J., Chang,J.,
Choepl,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Fero,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
```

Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.

#### TITLE

Submitted (10-APR-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

#### REFERENCE

4 (bases 1 to 196674)

#### AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,  
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,  
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,  
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,  
Erickson,J., Fero,S., Ferreira,P., FitzGerald,M., Gage,D.,  
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,  
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,  
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,  
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,  
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,  
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,  
Zimmer,A. and Zody,M.

#### TITLE

Submitted (01-MAY-2004) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

#### COMMENT

On May 1, 2004 this sequence version replaced gi:4638225.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submission@broad.mit.edu](mailto:sequence_submission@broad.mit.edu)

----- Project Information

Center project name: L18071

Center clone name: 202\_A\_19

#### FEATURES

##### source

Location/Qualifiers  
1..196674  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clone="RP23-202A19"  
/clone\_libs="RPMI-23 Female Mouse BAC"  
1..4229  
/note="wgs end extension  
clone end:SP6"  
repeat\_region  
1..193  
/rpt\_family="MTA"  
614..746  
/note="single clone coverage"  
701..706  
/note="<30 qual single clone coverage"  
1208..1291  
/rpt\_family="CTA)n"  
1571..1842  
/rpt\_family="BLTRETN MM"  
complement(1974..2067)  
/rpt\_family="B4"  
3219..3319  
/rpt\_family="GA-rich"

repeat\_region  
3368..3448  
/rpt\_family="CA)n"  
3626..3688  
/rpt\_family="CAAA)n"  
3925..4374  
/rpt\_family="Lx"  
4230..4235  
/note="clone boundary  
clone end:SP6  
site:ECORI  
4383..4411  
/rpt\_family="AT-rich"  
4445..4496  
/rpt\_family="TCTCTG)n"  
4509..4638  
/rpt\_family="TG)n"  
4741..4784  
/rpt\_family="TG)n"  
4785..4832  
/rpt\_family="TA)n"  
4833..4955  
/rpt\_family="TG)n"  
5398..5431  
/rpt\_family="CAAA)n"  
complement(5434..5561)  
/rpt\_family="RMER15"  
complement(6185..6272)  
/rpt\_family="RSINE1"  
6516..6570  
/rpt\_family="T)n"  
complement(6571..6823)  
/rpt\_family="L1\_MM"  
6821..6939  
/rpt\_family="L1\_MM"  
7007..7285  
/rpt\_family="Lx5"  
7292..7848  
/rpt\_family="Lx5"  
8187..8286  
/rpt\_family="RSINE1"  
complement(8823..9061)  
/rpt\_family="B4A"  
complement(9243..9815)  
/rpt\_family="L1"  
9832..9861  
/rpt\_family="A)n"  
9862..10368  
/rpt\_family="Lx2"  
complement(10373..11559)  
/rpt\_family="L1\_MM"  
11555..12196  
/rpt\_family="Lx2"  
complement(13221..13381)  
/rpt\_family="B3A"  
13437..13578  
/rpt\_family="B1\_MM"  
15222..15254  
/rpt\_family="AT-rich"  
complement(15315..15605)  
/rpt\_family="RMER20"  
complement(16008..16517)  
/rpt\_family="RMER5"  
16673..16713  
/rpt\_family="TG)n"  
16809..16866  
/rpt\_family="TG)n"  
complement(16922..17052)  
/rpt\_family="B4"  
17101..17490  
/rpt\_family="MLT2B1"  
17725..17762  
/rpt\_family="GGAA)n"

Query Match		46.0%; Score 706.8; DB 10; Length 196674;
Best Local Similarity		73.3%; Pred. No. 4.2e-173;
Matches 1013; Conservative		0; Mismatches 357; Indels 12; Gaps 8;
QY	143	TGTCGAGCTGGCCATTTCTGTGATGGAACAGCGGAGCAAGTGATCCCCCTGTGTGG 202
DB	101632	TTTATGACTATGCCATTTATGATGAAGAAAAGAAAGGATGCCCTTGTGTAG 101691
QY	203	GGGATGGACAGCTGCTCTCTAGAGATTGTAACTGGAGGAACCAACAGAGACTCTCAA 262
DB	101692	AGGATATAAGCTGCTCTCTGTAGGTGGCCACTTGTAGAAATCCTATGAGAATACTAA 101751
QY	263	ATACAGTTTGTATGCCCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATCAGGTTT 322
DB	101752	ATATCAGTTTAAATGCTCTTTATTTAGAGAAAGAGCACTCAGCAGAACATCTAAGTAGGTT 101811
QY	323	AGTCACGCCATCTTAAGGCACTTGTCTGAAGGAGGAGCAATGCCAGTGTGCTTCCCTATGG 382
DB	101812	AGTACGCCATCTCTGTGGCACTTGTCTGGATGGGAGATACCAAGTGTGCTTACATATGG 101871
QY	383	ACCACAGGAGCTGGGAGAGCGACACAATGCTGGGAGGCC----AGACAACTGGGG 438
DB	101872	GACTATTGAGCGAGGAGACACATAATACTGGGTAGCCCATGGGAAACAACATGGAG 101931
QY	439	TGATCCCGGGCTCTCATGAGACCTCTGCACTCACAAGGAGAGGGTGCCGAGGCC 498
DB	101932	TGATTCTTATGCTCTCATAGACCTCTGAGATTAACAAGGAGGAGGAGGCTGCTGAAGGCT 101991
QY	499	GGCCATGGCCCTTCTGTACCACTGCTTACCTAGAGATCTACCAAGGAGAGGTATTAG 558
DB	101992	GGCCATGGGACATTTCTGTAGGATGCTCTTATTAGAGTCTACCAAGGAGAGGTATTAG 102051
QY	559	ACCTCTGGACCTCTGCTTGGGAGACCTGTGTAATCCAGAGAGACTGCGGGGGGAATACC 618
DB	102052	ACCTCTTATATCTGCTCAGAGACCTAGTGTATGAGAAGATTGTGAGGGAACATCC 102111
QY	619	TGATTCCGGTCTCTCCAGAGGCCCATCAGTACTTGTGATTTGAGGGGCACTTCC 678
DB	102112	TGATTCCAGTTCTCACCAGGACCCCAATCTGCTTCTCTGACTTTGAGCAGTACTT-C 102170
QY	679	TGCCAGCGAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACAGCGCTCTCTCCGCA 738
DB	102171	TGCCAGCGCTGTAGAAATCAAGTTGTAGGAGTACTCTGCTTAAAGCAGTCTCTTACCATA 102230
QY	739	GTATGCTGTGCTCTCTGCTGAAGTGTGACAGCGGGAAGTGTGGCCCACTTTCGCCAGC 798
DB	102231	-TCATTCAATGCTCTTTATCAAGGTGGATCAGATGAGCAATTTGACTCTGTTTCACTAGT 102289
QY	799	GAGAGGGAATCTTACTGATTGACTTTGGCTGGGTGAGAGCAACCGGCGCACAGGCA 858
DB	102290	GAAAGGGAATCTTACCAATTTGATTTGGTTGGCTCAGAGGACCAACACACA-AGGGA 102348
QY	859	ACAAGGGCTTTGGCTTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGTCTGGGCA 918
DB	102349	ACCAGTGATTTCAACTCAAGAGAGTGGCGCATCAACACCTGCTTC-TTGTACTGGGCA 102407
QY	919	AAGTGTAGATGCGGTGATCAGGGCTCCCTCTGTGACTTATCGGAGCAGCAAGCTCA 978
DB	102408	AAGTGGAGTTGCAATATTAGGGCTCCCTCTCTATATTACTGGGACAGAAAAATCA 102467
QY	979	CTGGCTATTGAGGACTCTCTGGGTGGCTCAGGCCACAGTATCTTATTGCCAACATTG 1038
DB	102468	TTTGGCTATTGCGAGACTCTCTGGNAGGCTCAGCTCATAACATCATATTACCAACTG 102527
QY	1039	CCCTGAGAGAGCTTCTTACTAGACACAGTCTCGGCACTCAACTTTGCTGCCAGGTCCA 1098
DB	102528	CTCCTGAAAGACATTTTACTTGAATCATTTGAGCAATTTGAGCAATTTCACTGCTAGGTCCA 102587
QY	1099	AGGAGTCAATCGGCTTTTACCAATGAGCTGCGCTCATGCTCATGCTGGGACCTG 1158
DB	102588	AGGAGCTGATTAATTTGGGCTTTTACCAGTGGAGTCTACAGCTCATGCTTACACCTG 102647
QY	1159	TTAAGCTGTCTCAGAAAGAAATTTGCTTGTCCACCAGGAGCAAGAGAGCCCGAGGCCCTG 1218

Search completed: November 10, 2004, 13:18:01  
Job time : 6803.66 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2004, 09:45:10 ; Search time 795.24 Seconds  
(without alignments)  
10152.424 Million cell updates/sec

Title: US-10-797-893-3  
Perfect score: 1538  
Sequence: 1 atgcagccgcggcggtctg.....accattgtccacaatgtga 1538

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1538	100.0	1538	10	Adc23339 DNA encod
2	1538	100.0	1538	12	Adq60229 Human mic
3	1530	99.5	2097	12	Adq09241 Human KNS
4	1519	98.8	2099	10	Adk40995 Novel hum
5	1461	95.0	1464	10	Adc23343 DNA encod
6	1461	95.0	1464	12	Adq60233 Human mic
7	1450.2	94.3	2104	3	Aar15853 Human pro
8	1112.4	72.3	1115	10	Adc23337 DNA encod
9	1112.4	72.3	1115	12	Adq60227 Human mic
10	1035.4	67.3	1041	10	Adc23341 DNA encod
11	1035.4	67.3	1041	12	Adq60231 Human mic
12	480	31.2	491	6	Abk70260 Human lun
13	376.4	24.5	386	4	Aas38879 Novel hum
14	265	17.2	464	9	Ach44493 Human toe
15	233.2	15.2	531	12	Ach77408 Human gen
16	230.4	15.0	232	12	Ach91108 Human gen
17	185.6	12.1	2095	11	Adm01564 Human CDN
18	185.6	12.1	3624	12	Ado44181 DNA encod
19	156.8	10.2	3374	10	Adj95073 Novel NOV
20	154.8	10.1	2375	8	Abx34596 Human mdd
21	138.6	9.0	4108	6	Abx94614 Human kin

22	138.6	9.0	4108	8	ABS57215	Abx57215 Human kin
23	137.2	8.9	2675	10	ADJ95077	Adj95077 Novel NOV
24	136.6	8.9	1152	6	ABA94616	Abx94616 Nucleotid
25	136.6	8.9	1152	8	ABS57216	Abx57216 DNA encod
26	128.8	8.4	897	4	AAL70088	Aal70088 Human kin
27	128.8	8.4	897	6	AAL39619	Aal39619 DNA encod
28	128.8	8.4	897	12	ADM81051	Adm81051 Human HsK
29	127.8	8.3	488	5	ABV47991	Abv47991 Human pro
30	127.6	8.3	482	10	ABT40796	Abt40796 Toxicity
31	126.6	8.2	1026	6	ABQ73061	Abq73061 Human kin
32	126.6	8.2	1026	9	AAL56806	Aal56806 DNA encod
33	126.6	8.2	1026	10	ADB66785	Adb66785 DNA encod
34	124.2	8.1	1011	6	ABQ73062	Abq73062 Human kin
35	124.2	8.1	1011	9	AAL56807	Aal56807 DNA encod
36	124.2	8.1	1011	10	ADB66787	Adb66787 DNA encod
37	121.8	7.9	3661	6	AAD28568	Aad28568 Human kin
38	121.8	7.9	3694	10	ADC10189	Adc10189 Human NOV
39	121.4	7.9	1014	6	ABA94615	Abx94615 Human HsK
40	121.2	7.9	3570	12	ADQ19012	Adq19012 Human sof
41	119.8	7.8	2994	4	ABL28591	Abi28591 Drosophil
42	119.4	7.8	2034	4	ABL16733	Abi16733 Drosophil
43	119.4	7.8	4034	4	ABL16732	Abi16732 Drosophil
44	119.2	7.8	2132	10	ADC30338	Adc30338 Human nov
45	117	7.6	1839	10	ADC30569	Adc30569 Human nov

ALIGNMENTS

RESULT 1  
ADc23339  
ID ADc23339 standard; DNA; 1538 BP.  
XX  
AC ADc23339;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 3).  
XX  
KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
KW cardiac hypertrophy; immune disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 1..1538  
FT /\*tag= a  
FT /product= "KID protein"  
FT /transl\_except= (pos: 4..5; aa: Pro)  
FT /note= "This codon has an apparent 1 nucleotide deletion  
that alters the reading frame"

US6387644-B1.

14-MAY-2002.

28-NOV-2000; 2000US-00724224.

20-APR-1999; 99US-00295612.

20-JUN-2000; 2000US-00597292.

(CYTO-) CYTOKINETICS INC.

Beraud C;

WPI; 2003-706919/67.

P-PSDB; ADc23340.

Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that

directly/indirectly produces ADP or phosphate.

Disclosure; SEQ ID NO 3; 26pp; English.

This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cytosolic, cardants, immunomodulators and antiinflammatory. Accordingly, through gene therapy, they can be used for the treatment of cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders and inflammation. This polynucleotide sequence is human KID DNA (Seqid 3) encoding a full length KID enzyme of the invention.

Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;

Query Match 100.0%; Score 1538; DB 10; Length 1538;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCAGCGCGGGGGCTCGACGACGAGGCGACGCGAGATGGCGGAGCTTCACGG	60
Db	1	ATGCAGCGCGGGGGCTCGACGACGAGGCGACGCGAGATGGCGGAGCTTCACGG	60
Qy	61	CGGCGATCTCAGGAGCTGGTGGTGGTGAAGAGATGGAGTACTCGTCTCCAC	120
Db	61	CGGCGATCTCAGGAGCTGGTGGTGGTGAAGAGATGGAGTACTCGTCTCCAC	120
Qy	121	CTCCAGCTCGCGTAAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	180
Db	121	CTCCAGCTCGCGTAAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	180
Qy	181	CAAGTGATCCCCCTGGTGGGCGATGGACAGCTCTCTAGAGATGGTAACTGGA	240
Db	181	CAAGTGATCCCCCTGGTGGGCGATGGACAGCTCTCTAGAGATGGTAACTGGA	240
Qy	241	GGAAACACAGGAGACTCTCAATACCAATGATGCTTCTATGGGAGAGAGTACTC	300
Db	241	GGAAACACAGGAGACTCTCAATACCAATGATGCTTCTATGGGAGAGAGTACTC	300
Qy	301	AGCAGGACATCTATGAGGTTAGTGAGGCGGATCTTAAGGCACTTGGTGGAGGGG	360
Db	301	AGCAGGACATCTATGAGGTTAGTGAGGCGGATCTTAAGGCACTTGGTGGAGGGG	360
Qy	361	ATGCCAGTGTGCTTATGAGGCGGATGATCCCGGGGCTCTCATGGACCTCTGAGG	420
Db	361	ATGCCAGTGTGCTTATGAGGCGGATGATCCCGGGGCTCTCATGGACCTCTGAGG	420
Qy	421	GCCAGAGCAACCTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	480
Db	421	GCCAGAGCAACCTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	480
Qy	481	AGGAGGTGGCGAGGGCGGCGGATGGGCGGCTTCTGTCACATGTCTTACCTAGAT	540
Db	481	AGGAGGTGGCGAGGGCGGCGGATGGGCGGCTTCTGTCACATGTCTTACCTAGAT	540
Qy	541	ACCAGGAGAGGATATAGACTCTCTGACCTGCTTCGGGAGACCTGGTAACTCGGA	600
Db	541	ACCAGGAGAGGATATAGACTCTCTGACCTGCTTCGGGAGACCTGGTAACTCGGA	600
Qy	601	ACTGCCGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTG	660
Db	601	ACTGCCGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTG	660
Qy	661	ATTTTGAGCGGCACTTCTCCAGCGGATGGAATCGGATGAGGACCAACCGGCTCA	720
Db	661	ATTTTGAGCGGCACTTCTCCAGCGGATGGAATCGGATGAGGACCAACCGGCTCA	720

Qy	721	ACCAGCGCTCTCCGCGAGTCATGCTGTGCTCTCTGTCAGGTGGACCGGGAACGTT	780
Db	721	ACCAGCGCTCTCTCCGCGAGTCATGCTGTGCTCTCTGTCAGGTGGACCGGGAACGTT	780
Qy	781	TGGCCCCATTTTCCGACGAGAGGGAATACTCTACCTGATTGCTTGGCTGGGTACAGG	840
Db	781	TGGCCCCATTTTCCGACGAGAGGGAATACTCTACCTGATTGCTTGGCTGGGTACAGG	840
Qy	841	ACAACCGCGCACAGGCAACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCT	900
Db	841	ACAACCGCGCACAGGCAACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCT	900
Qy	901	CCCTGTTTGTCTGGGCAAGTGGTAGATCGCTGATCAGGGCTCCCTCGTGTACTTT	960
Db	901	CCCTGTTTGTCTGGGCAAGTGGTAGATCGCTGATCAGGGCTCCCTCGTGTACTTT	960
Qy	961	ATCGGACACCAAGCTCCTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCCACAGTA	1020
Db	961	ATCGGACACCAAGCTCCTCGCTATTGTCAGGACTCTCTGGGTGGCTCAGCCACAGTA	1020
Qy	1021	TCCTTATTGCAACATTTGCCCTTGAGAGAGCTTTCTACCTAGACACAGTCTCCGCACTCA	1080
Db	1021	TCCTTATTGCAACATTTGCCCTTGAGAGAGCTTTCTACCTAGACACAGTCTCCGCACTCA	1080
Qy	1081	ACTTGTGTCAGGTCGAGGAGTGAATCAATCGGCTTTTACCAATGAGAGCTCGAGC	1140
Db	1081	ACTTGTGTCAGGTCGAGGAGTGAATCAATCGGCTTTTACCAATGAGAGCTCGAGC	1140
Qy	1141	CTCATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGAGCAA	1200
Db	1141	CTCATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTTGTTGGTCCACAGAGCAA	1200
Qy	1201	AGAGAGCCGAGGCGCTTGAGGAGAGAGAGATGGGAGCCCTGAGCCCATGGAGTCCAG	1260
Db	1201	AGAGAGCCGAGGCGCTTGAGGAGAGAGAGATGGGAGCCCTGAGCCCATGGAGTCCAG	1260
Qy	1261	CCCTGCTCCCAAGAACTCAGCCCTCAGAAAGCTTAAGCAGCATGGACCCGCGCATGC	1320
Db	1261	CCCTGCTCCCAAGAACTCAGCCCTCAGAAAGCTTAAGCAGCATGGACCCGCGCATGC	1320
Qy	1321	TGAGAGCGCTCTCAGCTTGAGCCGCTGCTGCTCCCTCCAGGGGAGCAGGGGCGCCCTC	1380
Db	1321	TGAGAGCGCTCTCAGCTTGAGCCGCTGCTGCTCCCTCCAGGGGAGCAGGGGCGCCCTC	1380
Qy	1381	TGTTGATACCCCAAGACGAGAGCGGATGGTGTCTAATGAAGAAGTGAAGAAGAGACC	1440
Db	1381	TGTTGATACCCCAAGACGAGAGCGGATGGTGTCTAATGAAGAAGTGAAGAAGAGACC	1440
Qy	1441	TAGAGATTGAGAGGCTTAAAGACGAAGCAAAAGAACTGGAGGCAAGATGTTGGCCCGA	1500
Db	1441	TAGAGATTGAGAGGCTTAAAGACGAAGCAAAAGAACTGGAGGCAAGATGTTGGCCCGA	1500
Qy	1501	AGGCTGAGGAAAGAGGAGAACCATTTGCCCAATGTGA 1538	
Db	1501	AGGCTGAGGAAAGAGGAGAACCATTTGCCCAATGTGA 1538	

RESULT 2

ADQ60229

ID ADQ60229 standard; DNA; 1538 BP.

XX AC ADQ60229;

XX DT 23-SEP-2004 (first entry)

XX Human microtubule motor protein DNA #2.

XX Human; microtubule motor protein; gene; ds;

XX cellular proliferation disorder; cancer; hyperplasia; restenosis;

XX cardiac hypertrophy; immune disorder; inflammation;

XX kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;

XX KW graft rejection; inflammatory bowel disease; angioplasty.



OS Homo sapiens.  
 XX US6762043-B1.  
 XX 13-JUL-2004.  
 XX 06-MAR-2002; 2002US-00093317.  
 XX 20-APR-1999; 99US-00295612.  
 XX 20-JUN-2000; 2000US-00597292.  
 XX 28-NOV-2000; 2000US-00724224.  
 XX (CYTO-) CYTOKINETICS INC.  
 XX Beraud C;  
 XX WPI; 2004-532491/51.  
 XX New isolated microtubule motor protein, useful for screening modulators  
 XX for treating cellular proliferation disorders such as cancer,  
 XX hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
 XX inflammation.  
 XX Disclosure; SEQ ID NO 3; 26pp; English.  
 XX The invention relates to human microtubule motor proteins and the nucleic  
 XX acids encoding them. The invention also relates to a method of screening  
 XX for modulators of a motor protein which has microtubule stimulated ATPase  
 XX activity, a method of testing for ATPase activity of microtubule motor  
 XX protein, methods to identify candidate agents that bind to a target  
 XX protein or act as a modulator of the binding characteristics or  
 XX biological activity of a target protein, modulators of the target  
 XX protein, and methods of treating cellular proliferation disorders such as  
 XX cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
 XX and inflammation, for treating disorders associated with kinesin-like DNA  
 XX binding protein (KID) and for inhibiting KID. The sequences are used for  
 XX screening for modulators of motor proteins useful for treating cellular  
 XX proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
 XX hypertrophy, immune disorders and inflammation, for treating disorders  
 XX associated with KID and for inhibiting KID and for treating autoimmune  
 XX diseases, arthritis, graft rejection, inflammatory bowel disease and  
 XX proliferation induced after medical procedures including surgery and  
 XX angioplasty. This sequence represents DNA encoding a human microtubule  
 XX motor protein of the invention. Note: The specification states that this  
 XX sequence encodes the protein featured as SEQ ID NO:4, but this does not  
 XX appear to be the case.  
 XX Sequence 1538 BP; 355 A; 432 C; 459 G; 292 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 1538; DB 12; Length 1538;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGCAGCGCGCGGCTCGACGACGACGAGCGGCGGAGATGCGGCGAGCTTCAGCGG 60  
 DB 1 ATGCAGCGCGCGGCTCGACGACGACGAGCGGCGGAGATGCGGCGAGCTTCAGCGG 60  
 QY 61 CGGCGATCTCAGGAGCTGCTGCTCGCTAAGCAAGATTGGAGCTACTCGTCTCCAC 120  
 DB 61 CGGCGATCTCAGGAGCTGCTGCTCGCTAAGCAAGATTGGAGCTACTCGTCTCCAC 120  
 QY 121 CTCAGCTCGCTAAGGTTGGCTGTCGACTCGCGCCATTTGTGATGGAACAGCGGGAG 180  
 DB 121 CTCAGCTCGCTAAGGTTGGCTGTCGACTCGCGCCATTTGTGATGGAACAGCGGGAG 180  
 QY 181 CAAGTGATCCCCCTGTGTGCGGGGATGGAAGCTGCTCTAGAGATTGCTAACTGGA 240  
 DB 181 CAAGTGATCCCCCTGTGTGCGGGGATGGAAGCTGCTCTAGAGATTGCTAACTGGA 240  
 QY 241 GGAACACAGAGACTCTCAAAATACAGTTTGTGATGCTCTATGGGAGAGGAGTACTC 300  
 DB 241 GGAACACAGAGACTCTCAAAATACAGTTTGTGATGCTCTATGGGAGAGGAGTACTC 300

QY 301 AGCAGGACATCTATGTCAGGTTTCAGTGCAGCCCATCTTAAGGACACTTCTGCGAAGGGCAGA 360  
 DB 301 AGCAGGACATCTATGTCAGGTTTCAGTGCAGCCCATCTTAAGGACACTTCTGCGAAGGGCAGA 360  
 QY 361 ATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAGAGCACAATCTCTGGGCA 420  
 DB 361 ATGCCAGTGTGCTTGCCTATGACCCACAGGAGCTGGGAGAGCACAATCTCTGGGCA 420  
 QY 421 GCCCAGAGCAACTCGGGGTGATCCCGGGGCTCTCATGACCTCTCTGCAAGGG 480  
 DB 421 GCCCAGAGCAACTCGGGGTGATCCCGGGGCTCTCATGACCTCTCTGCAAGGG 480  
 QY 481 AGGAGGGTCCGAGGCGCGCCCATGGGCGCTTTCTGTCAACCATGTCTTACCTAGAGATCT 540  
 DB 481 AGGAGGGTCCGAGGCGCGCCCATGGGCGCTTTCTGTCAACCATGTCTTACCTAGAGATCT 540  
 QY 541 ACCAGAGAAAGTATTAGACCTCTTGGACCTCTGCTTGGGAGACCTGTGTAATCCGAGAAG 600  
 DB 541 ACCAGAGAAAGTATTAGACCTCTTGGACCTCTGCTTGGGAGACCTGTGTAATCCGAGAAG 600  
 QY 601 ACTGCGGGGGGATATCTCTGATTCGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTG 660  
 DB 601 ACTGCGGGGGGATATCTCTGATTCGGGTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTG 660  
 QY 661 ATTTTGGAGCGGCACTTCTCTGCGAGCAGTCGAAATCGGACTCTAGGAGCACCCGGCTCA 720  
 DB 661 ATTTTGGAGCGGCACTTCTCTGCGAGCAGTCGAAATCGGACTCTAGGAGCACCCGGCTCA 720  
 QY 721 ACCAGCGTCTCTCCGCACTATGTGCTCTCTGCTCAAGGTGGACCGGGAACGTT 780  
 DB 721 ACCAGCGTCTCTCCGCACTATGTGCTCTCTGCTCAAGGTGGACCGGGAACGTT 780  
 QY 781 TGGCCCCATTTCCGAGGAGGAGGAAACTCTACTGATGACTTGGTGGGTGAGAGG 840  
 DB 781 TGGCCCCATTTCCGAGGAGGAGGAAACTCTACTGATGACTTGGTGGGTGAGAGG 840  
 QY 841 ACACCGCGCGACAGCGCAACGAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCT 900  
 DB 841 ACACCGCGCGACAGCGCAACGAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCT 900  
 QY 901 CCTGTTTGTCTGGGCAAGTGTGATGCGCTGAATCAGGCGCTCTCCTGTGACCTT 960  
 DB 901 CCTGTTTGTCTGGGCAAGTGTGATGCGCTGAATCAGGCGCTCTCCTGTGACCTT 960  
 QY 961 ATCGGAGCAGCAAGCTCACTGCTGCTATGAGGACTCTCTGGTGGCTCAGCCCAAGTA 1020  
 DB 961 ATCGGAGCAGCAAGCTCACTGCTGCTATGAGGACTCTCTGGTGGCTCAGCCCAAGTA 1020  
 QY 1021 TCCTTTATTTGCCAACATTTGCCCTGAGAGACGCTTCTACTAGACAGTCTCCGCACTCA 1080  
 DB 1021 TCCTTTATTTGCCAACATTTGCCCTGAGAGACGCTTCTACTAGACAGTCTCCGCACTCA 1080  
 QY 1081 ACTTTGCTGCCAGGTCCAAAGGAGTGTCAATCGGCTTTTACCAATAGAGCCTGCGAGC 1140  
 DB 1081 ACTTTGCTGCCAGGTCCAAAGGAGTGTCAATCGGCTTTTACCAATAGAGCCTGCGAGC 1140  
 QY 1141 CTATGCTCTGGGACCTGTTAAGTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAA 1200  
 DB 1141 CTATGCTCTGGGACCTGTTAAGTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAA 1200  
 QY 1201 AGAGAGCCGAGGCGCTGAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAG 1260  
 DB 1201 AGAGAGCCGAGGCGCTGAGGAGAGGAGATTGGGAGCCCTGAGCCCATGGCAGCTCCAG 1260  
 QY 1261 CCTTGCCTCCAGAAACTCTAGCCCTTACAGAGCTAAGCAGCATGACCCGGCATGCG 1320  
 DB 1261 CCTTGCCTCCAGAAACTCTAGCCCTTACAGAGCTAAGCAGCATGACCCGGCATGCG 1320  
 QY 1321 TGGAGCGGCTCTCTCAGCTTTGGAACGCTGTGCTTGGCTCCAGGGGAGCAGGGGGCCCCTC 1380  
 DB 1321 TGGAGCGGCTCTCTCAGCTTTGGAACGCTGTGCTTGGCTCCAGGGGAGCAGGGGGCCCCTC 1380  
 QY 1381 TGTGTAGTACCCCAAGCGAGCGGATGCTGTCTAATGAAGCAGTAGAAGAGAGGAGC 1440





QY	126	GCTCGGTAAGGTTGCTGCGAATTCGCGCCATTTGTGATGGAACACGCGGAGCAAGT	185	QY	1206	GCCGAGGCCCTGAGAAAGAGAGATTGGAGCCCTGAGCCCATGCGAGCTCCAGCCTCT	1265
Db	149	GCTCGGTAAGGTTGCTGCGAATTCGCGCCATTTGTGATGGAACACGCGGAGCAAGT	208	Db	1229	GCCGAGGCCCTGAGAAAGAGAGATTGGAGCCCTGAGCCCATGCGAGCTCCAGCCTCT	1288
QY	186	GATCCCCCTGTTGCGGGCATGGACAGTGTCTCTAGAGATTCTTAACATGGAGAAC	245	QY	1266	GCCTCCCAAGAACTCAGCCCCCTTACAGAAAGCTTAAGCAGCATGGACCCCGCCATGCTGGAG	1325
Db	209	GATCCCCCTGTTGCGGGCATGGACAGTGTCTCTAGAGATTCTTAACATGGAGAAC	268	Db	1289	GCCTCCCAAGAACTCAGCCCCCTTACAGAAAGCTTAAGCAGCATGGACCCCGCCATGCTGGAG	1348
QY	246	CACACAGAGACTCTCAAAATACCAAGTTTGATGCCCTTATGGGAGAGAGTAGTACTCAGCAG	305	QY	1326	CGCCTCTCAGTTGGACCGTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTG	1385
Db	269	CACACAGAGACTCTCAAAATACCAAGTTTGATGCCCTTATGGGAGAGAGTAGTACTCAGCAG	328	Db	1349	CGCCTCTCAGTTGGACCGTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCTCTGTTG	1408
QY	306	GACATCTATGACGTTCAAGTGCAGCCCATCTTAAGGCACTTGTGTAAGGGGCAAGATGCC	365	QY	1386	AGTACCCCAAGCGAGAGCGGATGTCTTAATGAAGACAGTAGAAGAGAACCTAG-A	1444
Db	329	GACATCTATGACGTTCAAGTGCAGCCCATCTTAAGGCACTTGTGTAAGGGGCAAGATGCC	388	Db	1409	AGTACCCCAAGCGAGAGCGGATGTCTTAATGAAGACAGTAGAAGAGAACCTAGTA	1468
QY	366	AGTGTCTTCCCTATGACCCACAGAGCTGGGAAGCAGACACAATGCTGGGACGCCA	425	QY	1445	GATTGAGAGGCTTAAGACGAAGCAAAAGAACTGAGGCCCAAGATGTTGGCCCAAGAGGC	1504
Db	389	AGTGTCTTCCCTATGACCCACAGAGCTGGGAAGCAGACACAATGCTGGGACGCCA	448	Db	1469	GATTGAGAGGCTTAAGACGAAGCAAAAGAACTGAGGCCCAAGATGTTGGCCCAAGAGGC	1528
QY	426	GAGCAACTGGGTGATCCCGGGGCTCTCATGGACCTCTGAGCTCTCAAGGGAGGAG	485	QY	1505	TCAGAAAAGGAGAACCATTTGCCACAATG	1535
Db	449	GAGCAACTGGGTGATCCCGGGGCTCTCATGGACCTCTGAGCTCTCAAGGGAGGAG	508	Db	1529	TCAGAAAAGGAGAACCATTTGCCACAATG	1559
QY	486	GGTCCGAGGGCGGCGCATGGGCGCTTCTGTCAACATGTTTACCTAGAGATCTACCGAG	545	RESULT 5			
Db	509	GGTCCGAGGGCGGCGCATGGGCGCTTCTGTCAACATGTTTACCTAGAGATCTACCGAG	568	ID	ADC23343	standard; DNA; 1464 BP.	
QY	546	GAGAAAGTATTAGACCTCTCGGACCTGCTTCCGGAGACCTGTAATCCGAGAAAGACTGC	605	XX	ADC23343;		
Db	569	GAGAAAGTATTAGACCTCTCGGACCTGCTTCCGGAGACCTGTAATCCGAGAAAGACTGC	628	XX	18-DEC-2003	(first entry)	
QY	606	CGGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTT	665	DE	DNA	encoding the human kinesin-like DNA binding protein (KID) (SeqID 7).	
Db	629	CGGGGGAATATCTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTT	688	KW	human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;		
QY	666	GAGGGCACTTCTGCGCAGCAGTGCAGAAATCGACTGTAGAGCCACCGGCTCAACCCAG	725	KW	cytostatic; cardiant; immunomodulator; antiinflammatory; gene therapy;		
Db	689	GAGGGCACTTCTGCGCAGCAGTGCAGAAATCGACTGTAGAGCCACCGGCTCAACCCAG	748	KW	cancer; hyperplasia; restenosis; cellular proliferation disorder;		
QY	726	CGCTCTCCGCGAGTCTGCTGCTCTGCTCAAGGTGGACAGCGGGAACTGTTGGCC	785	XX	cardiac hypertrophy; immune disorder; inflammation.		
Db	749	CGCTCTCCGCGAGTCTGCTGCTCTGCTCAAGGTGGACAGCGGGAACTGTTGGCC	808	OS	Homo sapiens.		
QY	786	CCATTTCCGCGAGGAGGAAACTCTA CTTGATTCAGTTCGCTGGGTGAGAGCAAC	845	XX	Key	Location/Qualifiers	
Db	809	CCATTTCCGCGAGGAGGAAACTCTA CTTGATTCAGTTCGCTGGGTGAGAGCAAC	868	XX	1. .1464		
QY	846	CGGCGCACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTG	905	XX	/tag= a	/product= "KID protein"	
Db	869	CGGCGCACAGGCAACAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTG	928	XX			
QY	906	TTTGTCTGGGCAAGTGTAGATGCGTGAATCAGGGCTCCCTCGTGTACCTTATCGG	965	XX	US6387644-B1.		
Db	929	TTTGTCTGGGCAAGTGTAGATGCGTGAATCAGGGCTCCCTCGTGTACCTTATCGG	988	XX	14-MAY-2002.		
QY	966	GACAGCAAGCTCACTCGCCTATTGAGGAGTCTCTGGGTGGCTCAGCCCAAGTATCCTT	1025	XX	28-NOV-2000; 2000US-00724224.		
Db	989	GACAGCAAGCTCACTCGCCTATTGAGGAGTCTCTGGGTGGCTCAGCCCAAGTATCCTT	1048	XX	20-APR-1999; 99US-00295612.		
QY	1026	ATTGCCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT	1085	XX	20-JUN-2000; 2000US-00597292.		
Db	1049	ATTGCCAATTTGCCCTGAGAGAGCTTCTACCTAGACACAGTCTCCGCACTCAACTTT	1108	XX	(CYTO-) CYTOKINETICS INC.		
QY	1086	GCTGCGAGTCCAGGAGGTGATCAATCGGCTTTTACCATGAGAGCCTGCAGCCTCAT	1145	XX	Beraud C;		
Db	1109	GCTGCGAGTCCAGGAGGTGATCAATCGGCTTTTACCATGAGAGCCTGCAGCCTCAT	1168	XX	WPI; 2003-706919/67.		
QY	1146	GCCTTGGAGACCTGTTAGCTGCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGA	1205	XX	P-PSDB; ADC23344.		
Db	1169	GCCTTGGAGACCTGTTAGCTGCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGA	1228	XX	Identifying a candidate agent as modulator of function of a target protein for treating cellular proliferation disorders by adding a candidate agent to a mixture of the target protein that directly/indirectly produces ADP or phosphate.		
				XX	Disclosure; SEQ ID NO 7; 26pp; English.		
				CC	This invention relates to a novel method for high throughput screening systems used to identify compounds for the treatment of cellular proliferation disorders. Specifically, it refers to candidate agents that		

are capable of modulating the activity of target proteins having motor domains, such that the target protein directly or indirectly produces ADP or phosphate. Furthermore, this activity can be determined using fluorescence or absorbance readouts. The present invention describes a method that identifies modulators of the target protein, which is a kinesin-like DNA binding protein (known as KID) as cytoskeletal, cardiac, immunomodulators and antiinflammatory. Accordingly, through gene therapy, they can be used for the treatment of cancer, hyperplasia, stenosis, cardiac hypertrophy, immune disorders and inflammation. This polynucleotide sequence is human KID DNA (Seqid 7) encoding a full length KID enzyme of the invention.

**SQ** Sequence 1464 BP; 342 A; 409 C; 428 G; 285 T; 0 U; 0 Other;

```
Query Match      95.0%; Score 1461; DB 10; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1461: Conservative 0; Mismatches 0; Indels 0;
```

Qy	78	GGT	CGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACTCCAGCTCGCGTAAGG	137
Db	4	GGT	CGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTGTCCACTCCAGCTCGCGTAAGG	63
Qy	138	GTGGCTGTGCGACTCGGCGCAATTTGTGATCGGAACAGCGGAGCAAGTGTATCCCCCTGT	197	
Db	64	GTGGCTGTGCGACTCGGCGCAATTTGTGATCGGAACAGCGGAGCAAGTGTATCCCCCTGT	123	
Qy	198	GTGCGGGCAITGGACAGCTGCTCTCTAGAGATTGCTAACTTGGAGGAACACACAGGAGACT	257	
Db	124	GTGCGGGCAITGGACAGCTGCTCTCTAGAGATTGCTAACTTGGAGGAACACACAGGAGACT	183	
Qy	258	CTCAAAATACCAGTTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCMAGGACATCTATGCA	317	
Db	184	CTCAAAATACCAGTTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA	243	
Qy	318	GGTTCAGTGCACGCCATCCTAAGGCACCTTGCTGGAAGGCGAGAAATGCCAGTGTGCTTGCC	377	
Db	244	GGTTCAGTGCACGCCATCCTAAGGCACCTTGCTGGAAGGCGAGAAATGCCAGTGTGCTTGCC	303	
Qy	378	TATGGAACCCACAGGAGCTGGGAAGACGCACA CAATGCTGGGCGAGCCACAGACCAACTGGG	437	
Db	304	TATGGAACCCACAGGAGCTGGGAAGACGCACA CAATGCTGGGCGAGCCACAGACCAACTGGG	363	
Qy	438	GTGATCCCGCGGGCTTCATGAGCTCTCTGTCAGCTCAAGGGAGAGGGGTGCGGAGGGC	497	
Db	364	GTGATCCCGCGGGCTTCATGAGCTCTCTGTCAGCTCAAGGGAGAGGGGTGCGGAGGGC	423	
Qy	498	CGGCCATGGGGCCCTTTCTGTCAACATGTCTTACCCTAGAGATCTACACAGAGAGAGTATTA	557	
Db	424	CGGCCATGGGGCCCTTTCTGTCAACATGTCTTACCCTAGAGATCTACACAGAGAGAGTATTA	483	
Qy	558	GACCTCTTGGACCTTGCTTTCGGGAGACCTGTGTAAATCCGAGAGACTTGCCTGGGGGAAATTC	617	
Db	484	GACCTCTTGGACCTTGCTTTCGGGAGACCTGTGTAAATCCGAGAGACTTGCCTGGGGGAAATTC	543	
Qy	618	CTGATTTCGGGGTCTCTCCACAGAGCCCATCAGTAGCTTTTCTGATTTTGAGCGGCACTTC	677	
Db	544	CTGATTTCGGGGTCTCTCCACAGAGCCCATCAGTAGCTTTTCTGATTTTGAGCGGCACTTC	603	
Qy	678	CTGCCAGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGCGCTCTCTCCCGC	737	
Db	604	CTGCCAGCAGTCGAAATCGGACTGTAGGAGCCACCCGGCTCAACACAGCGCTCTCTCCCGC	663	
Qy	738	AGTCATCTGTGCTCTCGTCAAGGTGGACGAGGGGAAACGTTTGGCCCCCATTTGCCGAG	797	
Db	664	AGTCATCTGTGCTCTCGTCAAGGTGGACGAGGGGAAACGTTTGGCCCCCATTTGCCGAG	723	
Qy	798	CGAGAGGGAAACCTACCTGATTTGACTTGTGCTGGGTTCAGAGCAACAACCGGCGCACAGGC	857	
Db	724	CGAGAGGGAAACCTACCTGATTTGACTTGTGCTGGGTTCAGAGCAACAACCGGCGCACAGGC	783	
Qy	858	AACAAGGGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTGTTGTCTGGGC	917	
Db	784	AACAAGGGCTTCGGCTAAAAGAGAGTGGAGCCATCAACACCTCCCTGTGTTGTCTGGGC	843	



RESULT 7										
AAF15853										
ID	AAF15853 standard; cDNA; 2104 BP.									
XX										
AC	AAF15853;									
XX										
DT	13-MAR-2001 (first entry)									
XX										
XX	Human prostate cancer antigen nucleotide sequence SEQ ID NO:288.									
DE										
XX										
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;									
KW	neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;									
KW	vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;									
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;									
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;									
KW	wound; infectious disease; ss.									
XX										
OS	Homo sapiens.									
XX										
PN	W0200055174-A1.									
XX										
PD	21-SEP-2000.									
XX										
PP	08-MAR-2000; 2000WO-US005988.									
XX										
PR	12-MAR-1999; 99US-0124270P.									
XX										
PA	(HUMA-) HUMAN GENOME SCI INC.									
PA	(ROSE/) ROSEN C A.									
XX										
PI	Rosen CA, Ruben SM;									
XX										
DR	WPI; 2000-587513/55.									
DR	P-PSDB; AAB56650.									
XX										
PT	Prostate cancer associated gene sequences, referred to as prostate cancer									
PT	antigens, useful for treatment, prevention, and diagnosis of disorders									
PT	such as prostate cancer.									
XX										
PS	Claim 1; Page 805-806; 2338pp; English.									
XX										
CC	AAF15566 to AAF16505 encode the human prostate cancer associated									
CC	proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.									
CC	The prostate cancer antigens can have neuroprotective, cytostatic,									
CC	cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,									
CC	nephrotropic, antineoplastic, gynaecological and antibacterial activities,									
CC	and can be used in gene therapy. The prostate cancer antigen									
CC	polynucleotides may be used for detection of prostate cancer, chromosome									
CC	identification, as chromosome markers, and for numerous other diagnostic									
CC	or research purposes. The prostate cancer antigens may be used to treat									
CC	disorders such as neural, immune, muscular, reproductive,									
CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative									
CC	disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to									
CC	AAB57303 represent sequences used in the exemplification of the present									
XX	invention									
XX										
SQ	Sequence 2104 BP; 517 A; 580 C; 597 G; 400 T; 0 U; 10 Other;									
Query Match 94.3%; Score 1450.2; DB 3; Length 2104;										
Best Local Similarity 99.4%; Pred. No. 0;										
Matches 1457; Conservative 7; Mismatches 1; Indels 1; Gaps 1;										
Qy	71	AGGAGCTGGTGGCTGCGCTAAGCAGATTGGAGCTACTCGTCGTCACCTCCAGCTGC	130							
Db	60	AGGAGCTGGTGGCTGCGCTAAGCAGATTGGAGCTACTCGTCGTCACCTCCAGCTGC	119							
Qy	131	CGTAAGGGTGGCTGTCGACTCGGCCATTGTCGATGGAACAGCGGAGCAAGTGATCC	190							
Db	120	CGTAAGGGTGGCTGTCGACTCGGCCATTGTCGATGGAACAGCGGAGCAAGTGATCC	179							
Qy	191	CCCCTGTGTGGGGGCATGGAACAGCTGCTCTTAGAGATTGCTTAAGTGGAGAACACCA	250							









901 CCTGTTTGTCTGGGCAAGTGGTAGATGCGTGAATCAGGCGCTCCCTCGTGTACCTT 960  
Cc  
901 CCTGTTTGTCTGGGCAAGTGGTAGATGCGTGAATCAGGCGCTCCCTCGTGTACCTT 960  
Cc  
961 ATCGGACAGCAAGCTCAGCTCGCTATTGCGAGACTCTCTGGTGGCTCAGCCACAGTA 1020  
Cc  
961 ATCGGACAGCAAGCTCAGCTCGCTATTGCGAGACTCTCTGGTGGCTCAGCCACAGTA 1020  
Cc  
1021 TCCTATTGCAACATTGCCCCGAGAGCGCTTCTACCTAGACACAGCTCTCGCACTCA 1080  
Cc  
1021 TCCTATTGCAACATTGCCCCGAGAGCGCTTCTACCTAGACACAGCTCTCGCACTCA 1080  
Cc  
1081 ACTTTGTCGAGGTCGAGGAGGTGATCAATCG 1114  
Cc  
1081 ACTTTGTCGAGGTCGAGGAGGTGATCAATCG 1114  
Cc  
RESULT 10  
ID ADC23341 standard; DNA; 1041 BP.  
XX AC ADC23341;  
XX DT 18-DEC-2003 (first entry)  
XX DE DNA encoding the human kinesin-like DNA binding protein (KID) (SeqID 5).  
XX KW human; gene; ds; motor domain; kinesin-like DNA binding protein; KID;  
XX KW cytosolic; cardiant; immunomodulator; antiinflammatory; gene therapy;  
XX KW cancer; hyperplasia; restenosis; cellular proliferation disorder;  
XX KW cardiac hypertrophy; immune disorder; inflammation.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT 1..1041  
XX FT /\*tag= a  
XX FT /product= "KID protein"  
XX US6387644-B1.  
XX PD 14-MAY-2002.  
XX PF 28-NOV-2000; 2000US-00724224.  
XX PR 20-APR-1999; 99US-00295612.  
XX PR 20-JUN-2000; 2000US-00597292.  
XX PA (CYTO-) CYTOKINETICS INC.  
XX PI Beraud C;  
XX WPI; 2003-706919/67.  
XX DR P-PSDB; ADC23342.  
XX XX Identifying a candidate agent as modulator of function of a target  
XX PT protein for treating cellular proliferation disorders by adding a  
XX PT candidate agent to a mixture of the target protein that  
XX PT directly/indirectly produces ADP or phosphate.  
XX PS Disclosure; SEQ ID NO 5; 26pp; English.  
XX CC This invention relates to a novel method for high throughput screening  
XX CC systems used to identify compounds for the treatment of cellular  
XX CC proliferation disorders. Specifically, it refers to candidate agents that  
XX CC are capable of modulating the activity of target proteins having motor  
XX CC domains, such that the target protein directly or indirectly produces ADP  
XX CC or phosphate. Furthermore, this activity can be determined using  
XX CC fluorescence or absorbance readouts. The present invention describes a  
XX CC method that identifies modulators of the target protein, which is a  
XX CC kinesin-like DNA binding protein (known as KID) as cytosolic, cardiants,  
XX CC immunomodulators and antiinflammatories. Accordingly, through gene

therapy, they can be used for the treatment of cancer, hyperplasias,  
Cc restenosis, cardiac hypertrophy, immune disorders and inflammation. This  
Cc polynucleotide sequence is human KID DNA (SeqID 5) encoding a full length  
Cc KID enzyme of the invention.  
XX  
SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;  
Query Match 67.3%; Score 1035.4; DB 10; Length 1041;  
Best Local Similarity 99.9%; Pred. No. 1.1e-274;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 78 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTGCACCTCCAGCTCGCTAAGG 137  
Db 4 GGTGCTGTGCGCTAAGCAAGATTGGAGCTACTCGTCTGCACCTCCAGCTCGCTAAGG 63  
QY 138 GTGGCTGTGCGACTGCGGCCATTGTGGATGAACAGCGGGAGCAAGTGAATCCCCCTGT 197  
Db 64 GTGGCTGTGCGACTGCGGCCATTGTGGATGAACAGCGGGAGCAAGTGAATCCCCCTGT 123  
QY 198 GTGGCGGCGATGAGCAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGCT 257  
Db 124 GTGGCGGCGATGAGCAGCTGCTCTAGAGATTGCTAACTGGAGGAACACACAGGAGCT 183  
QY 258 CTCAAAATACCACTTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317  
Db 184 CTCAAAATACCACTTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243  
QY 318 GGTTCAGTCAGCCCATCTTAAGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGGC 377  
Db 244 GGTTCAGTCAGCCCATCTTAAGCACTTGTGGAGGGCAGAAATGCCAGTGTGCTTGGC 303  
QY 378 TATGAGCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCGCCACAGCAACCTGGG 437  
Db 304 TATGAGCCACAGGAGCTGGGAAGACGACACAAATGCTGGGAGCGCCACAGCAACCTGGG 363  
QY 438 GTGATCCCGCGGCTCTCATGAGCTCTGAGAGCTCAAGGAGGAGGGTCCGAGGGC 497  
Db 364 GTGATCCCGCGGCTCTCATGAGCTCTGAGAGCTCAAGGAGGAGGGTCCGAGGGC 423  
QY 498 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTATTA 557  
Db 424 CGGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTATTA 483  
QY 558 GACCTCTCTGAGACCTCTTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAATATC 617  
Db 484 GACCTCTCTGAGACCTCTTTCGGGAGACCTGGTAATCCGAGAGACTGCGGGGGAATATC 543  
QY 618 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 677  
Db 544 CTGATTCGGGCTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTTCAGCGGCACTTC 603  
QY 678 CTGCCAGCCAGTCGAAATCGGAGCTGAGGAGCCACCGGCTCAACAGCGCTCCTCCCGC 737  
Db 604 CTGCCAGCCAGTCGAAATCGGAGCTGAGGAGCCACCGGCTCAACAGCGCTCCTCCCGC 663  
QY 738 AGTCATGCTGTGCTCTCTGCTCAAGTGGACACAGCGGAAACGTTTGGCCCCCATTTTCGCGAG 797  
Db 664 AGTCATGCTGTGCTCTCTGCTCAAGTGGACACAGCGGAAACGTTTGGCCCCCATTTTCGCGAG 723  
QY 798 CGAGAGGAAACCTCTACCTGATTGCTTGGCTGAGGAGCAACCGGCGCACAGGC 857  
Db 724 CGAGAGGAAACCTCTACCTGATTGCTTGGCTGAGGAGCAACCGGCGCACAGGC 783  
QY 858 AACAGGGCTCTCGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 917  
Db 784 AACAGGGCTCTCGCTTAAAGAGAGTGGAGCCATCAACACCTCTCTGTTTGTCTGGGC 843  
QY 918 AAAGTGTAGATCGCTGAATAGGGCTCCCTCGTGTACCTTATCGGAGACAGAGCTC 977  
Db 844 AAAGTGTAGATCGCTGAATAGGGCTCCCTCGTGTACCTTATCGGAGACAGAGCTC 903  
QY 978 ACTCGCTATTGAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTATTTGCCAATT 1037  
Db 903 ACTCGCTATTGAGGAGCTCTCTGGTGGCTCAGCCACAGTATCTTATTTGCCAATT 1037

Db 904 ACTCGCTATTGACAGACTCTCTGGTGGCTCAGCCACAGATATCTTATTGCCAACATT 963  
Qy 1038 GCCCTCTGAGAGAGCGTTCTTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAGGTCC 1097  
Db 964 GCCCTCTGAGAGAGCGTTCTTACCTAGACACAGTCTCGCACTCAACTTTGCTGCCAGGTCC 1023  
Qy 1098 AAGGAGGTGATCAATCG 1114  
Db 1024 AAGGAGGTGATCAATG 1040

RESULT 11  
ADQ60231  
ID ADQ60231 standard; DNA; 1041 BP.  
XX  
AC ADQ60231;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Human microtubule motor protein DNA #3.  
XX  
KW Human; microtubule motor protein; gene; ds;  
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
KW cardiac hypertrophy; immune disorder; inflammation;  
KW kinesin-like DNA binding protein; KID; autoimmune disease; arthritis;  
KW graft rejection; inflammatory bowel disease; angioplasty.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1041  
FT FT /\*tag= a  
FT FT /product= "Human microtubule motor protein #3"  
XX  
PN US6762043-B1.  
XX  
PD 13-JUL-2004.  
XX  
XX 06-MAR-2002; 2002US-00093317.  
XX  
XX 20-APR-1999; 98US-00295612.  
XX 20-JUN-2000; 2000US-00597292.  
XX 28-NOV-2000; 2000US-00724224.  
XX  
XX (CYTO-) CYTOKINETICS INC.  
XX  
XX Beraud C;  
XX  
XX WPI; 2004-532491/51.  
XX P-PSDB; ADQ60232.  
XX  
XX New isolated microtubule motor protein, useful for screening modulators  
XX for treating cellular proliferation disorders such as cancer,  
XX hyperplasias, restenosis, cardiac hypertrophy, immune disorders and  
XX inflammation.  
XX  
XX Disclosure; SEQ ID NO 5; 26pp; English.  
XX  
XX The invention relates to human microtubule motor proteins and the nucleic  
XX acids encoding them. The invention also relates to a method of screening  
XX for modulators of a motor protein which has microtubule stimulated ATPase  
XX activity, a method of testing for ATPase activity of microtubule motor  
XX proteins, methods to identify candidate agents that bind to a target  
XX protein or act as a modulator of the binding characteristics or  
XX biological activity of a target protein, modulators of the target  
XX protein, and methods of treating cellular proliferation disorders such as  
XX cancer, hyperplasias, restenosis, cardiac hypertrophy, immune disorders  
XX and inflammation, for treating disorders associated with kinesin-like DNA  
XX binding protein (KID) and for inhibiting KID. The sequences are used for  
XX screening for modulators of motor proteins useful for treating cellular  
XX proliferation disorders such as cancer, hyperplasias, restenosis, cardiac  
XX hypertrophy, immune disorders and inflammation, for treating disorders  
XX associated with KID and for inhibiting KID and for treating autoimmune

CC diseases, arthritis, graft rejection, inflammatory bowel disease and  
CC proliferation induced after medical procedures including surgery and  
CC angioplasty. This sequence represents DNA encoding a human microtubule  
CC motor protein of the invention.  
XX  
SQ Sequence 1041 BP; 228 A; 295 C; 302 G; 216 T; 0 U; 0 Other;  
Query Match 67.3%; Score 1035.4; DB 12; Length 1041;  
Best Local Similarity 99.9%; Pred. No. 1.le-274;  
Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 78 GGTCTGCTCGGCTAAGCAAGATTGGAGCTACTCTGCTGCCACCTCCAGCTCCGCTAAGG 137  
Db 4 GGTCTGCTCGGCTAAGCAAGATTGGAGCTACTCTGCTGCCACCTCCAGCTCCGCTAAGG 63  
Qy 138 GTGGCTGTGCAGCTGCGGCCCATTTGTGATGGAAACAGCGGGAGCAAGTATCCCCCTGT 197  
Db 64 GTGGCTGTGCAGCTGCGGCCCATTTGTGATGGAAACAGCGGGAGCAAGTATCCCCCTGT 123  
Qy 198 GTGGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACGAGGACT 257  
Db 124 GTGGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACACACGAGGACT 183  
Qy 258 CTCAAAATACCAAGTTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 317  
Db 184 CTCAAAATACCAAGTTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACATCTATGCA 243  
Qy 318 GTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGNAGGGCAGATGCCAGTGTGCTGCC 377  
Db 244 GTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGNAGGGCAGATGCCAGTGTGCTGCC 303  
Qy 378 TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCAGAGCAACCTGGG 437  
Db 304 TATGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCAGAGCAACCTGGG 363  
Qy 438 GTGATCCCGGGCTCTCATGGACCTCTGTCAGCTCAAGAGGAGAGGGTGCAGAGGC 497  
Db 364 GTGATCCCGGGCTCTCATGGACCTCTGTCAGCTCAAGAGGAGAGGGTGCAGAGGC 423  
Qy 498 CGGCCATGGGCCCTTTCTGTCAACCATGCTTACCTAGAGATCTACCAGAGAGATATTA 557  
Db 424 CGGCCATGGGCCCTTTCTGTCAACCATGCTTACCTAGAGATCTACCAGAGAGATATTA 483  
Qy 558 GACCTCTGGACCTGCTTCCGGAGAGCTGTGTAATCCGAGAGAGCTGCCGGGGATATC 617  
Db 484 GACCTCTGGACCTGCTTCCGGAGAGCTGTGTAATCCGAGAGAGCTGCCGGGGATATC 543  
Qy 618 CTGATTCGGGTCTCTCCAGAACCCATCATAGTAGCTTTGTGATTTTGGCGGCACCTTC 677  
Db 544 CTGATTCGGGTCTCTCCAGAACCCATCATAGTAGCTTTGTGATTTTGGCGGCACCTTC 603  
Qy 678 CTGCCAGCCAGTCAAAATCGGACTGTAGGAGCCACCGGCTCAACAGGCTCTCCCGC 737  
Db 604 CTGCCAGCCAGTCAAAATCGGACTGTAGGAGCCACCGGCTCAACAGGCTCTCCCGC 663  
Qy 738 AGTCATGCTGTGCTCTGGTCAAGTGGACCGGGAACCTTTGGGCCCATTTTGCACAG 797  
Db 664 AGTCATGCTGTGCTCTGGTCAAGTGGACCGGGAACCTTTGGGCCCATTTTGCACAG 723  
Qy 798 CGAGAGGAAATCTCTACCTGATTTGACTTGGCTGGGTGAGAGCAACCCGCGCAGAGC 857  
Db 724 CGAGAGGAAATCTCTACCTGATTTGACTTGGCTGGGTGAGAGCAACCCGCGCAGAGC 783  
Qy 858 AACAGGGCTCTCGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCGGC 917  
Db 784 AACAGGGCTCTCGCTAAAGAGAGTGGAGCCATCAACACTCCCTGTTTGTCTCGGC 843  
Qy 918 AAAGTGTAGATGGCTGGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 977  
Db 844 AAAGTGTAGATGGCTGGAATCAGGGCTCCCTCGTGTACCTTATCGGAGCAGCAAGCTC 903  
Qy 978 ACTCGCTATTGACAGGACTCTCTGGGTGGGTTCAGGCCCAAGTATCTTATTTGCCAACATT 1037

Db 904 ACTCGCCTATTGAGGACTCTCTGGTGGCTGACGCCACAGTATCCTTTATTGCCAATT 963  
QY 1038 GCCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCGAGTCC 1097  
Db 964 GCCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCGAGTCC 1023  
QY 1098 AAGGAGGTGATCAATCG 1114  
Db 1024 AAGGAGGTGATCAATTG 1040

RESULT 12  
ABK70260  
ID ABK70260 standard; cDNA; 491 BP.  
AC ABK70260;  
XX  
XX  
DT 15-JUL-2002 (first entry)  
XX Human lung cancer associated cDNA SEQ ID 131.  
XX DE Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.  
XX KW Homo sapiens.  
XX OS  
XX PN WO200224057-A2.  
XX PD  
XX PF 28-MAR-2002.  
XX PP 20-SEP-2001; 2001WO-US042232.  
XX PR 22-SEP-2000; 2000US-0234837P.  
XX PR 10-OCT-2000; 2000US-0239440P.  
XX PR 29-JUN-2001; 2001US-0301928P.  
XX  
XX PA (CORI-) CORIXA CORP.  
XX Benson DR, Mohamath R, Lodes MJ;  
XX WPI; 2002-372001/40.  
XX  
XX New tumor lung proteins and nucleic acids encoding the proteins, useful  
PT as vaccines and for treating, preventing, diagnosing or monitoring lung  
PT cancer.  
XX  
XX Claim 1; Page 144; 189pp; English.

The invention relates to an isolated polynucleotide comprising a sequence  
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),  
CC or their fragments, homologues, variants or complements and their encoded  
CC polypeptides. Also included are an expression vector comprising the  
CC polynucleotide operably linked to an expression control sequence; a host  
CC cell transformed or transfected with an expression vector of; an isolated  
CC antibody, or its antigen-binding fragment that specifically binds to the  
CC polypeptide; a method for detecting the presence of a cancer in a patient  
CC; a fusion protein comprising at least the polypeptide; an  
CC oligonucleotide that hybridizes to the polynucleotide under moderately  
CC stringent conditions; a method for stimulating and/or expanding T cells  
CC specific for a tumour protein; an isolated T cell population comprising T  
CC cells prepared from the method of above; a composition comprising a first  
CC component consisting of carriers and immunostimulants, and a second  
CC component selected from the polynucleotides, proteins, antibodies, fusion  
CC proteins, T cell populations and antigen presenting cells expressing the  
CC polypeptide; methods for stimulating an immune response or treating  
CC cancer in a patient by administering the composition and diagnostic kits  
CC comprising at least one of the oligonucleotide of, or an antibody and a  
CC detection reagent consisting of a reporter group. The polypeptides and  
CC polynucleotides are useful as vaccines for the treatment or prevention of  
CC lung cancer, and for diagnosis and monitoring of such cancer. The  
CC polynucleotide, polypeptide and antigen presenting cells can be used to  
CC stimulate or expand T cells specific for a tumorous protein. The  
CC polynucleotides may be used as probes or primers for nucleic acid  
CC hybridisation, and in the preparation of ribozyme molecules for

CC inhibiting expression of tumour polypeptides and proteins in tumour  
CC cells. The present sequence is one of the 183 lung cancer associated  
CC polynucleotides  
XX  
SQ Sequence 491 BP; 127 A; 140 C; 136 G; 88 T; 0 U; 0 Other;  
Query Match 31.2%; Score 480; DB 6; Length 491;  
Best Local Similarity 99.8%; Pred. No. 9.3e-122;  
Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1010 AGCCCAAGTATCCTTTATTGCCAACATTTGCCCTGAGAGCGCTTCTTACCTAGACACAGT 1069  
Db 1 AGCCCAAGTATCCTTTATTGCCAACATTTGCCCTGAGAGCGCTTCTTACCTAGACACAGT 60  
QY 1070 CTCGGCACTCAACTTTGCTGCCAGGTCCCAAGAGGTGATCAATCGGCTTTTACCAATGA 1129  
Db 61 CTCGGCACTCAACTTTGCTGCCAGGTCCCAAGAGGTGATCAATCGGCTTTT-CCAATGA 1119  
QY 1130 GAGCCTGCGAGCCTCATGCCCTTGGGACCTGTTAAGCTGTCTCAGAAAGAATTGCTTGGTCC 1189  
Db 120 GAGCCTGCGAGCCTCATGCCCTTGGGACCTGTTAAGCTGTCTCAGAAAGAATTGCTTGGTCC 1179  
QY 1190 ACCAGAGGCAAGAGAGAGCCCGAGGCCCTGAGGAAGAGGAGATTGGGAGCCCTGAGCCCAT 1249  
Db 180 ACCAGAGGCAAGAGAGAGCCCGAGGCCCTGAGGAAGAGGAGATTGGGAGCCCTGAGCCCAT 239  
QY 1250 GGCAGCTCAGCCTCTGCCCTCCAGAAACTCAGCCCCCTTACAGAAGCTAAGAGCATGGA 1309  
Db 240 GGCAGCTCAGCCTCTGCCCTCCAGAAACTCAGCCCCCTTACAGAAGCTAAGAGCATGGA 299  
QY 1310 CCGGGCCATGCTGGAGCGCCTCCTCAGCTTGGACCGCTGCTTGGCTCCCGAGGGAGCCA 1369  
Db 300 CCGGGCCATGCTGGAGCGCCTCCTCAGCTTGGACCGCTGCTTGGCTCCCGAGGGAGCCA 359  
QY 1370 GGGGGCCCTCTGTTGAGTACCCCAAGCGAGAGCGGATGGTCTTAATGAAGACAGTAGA 1429  
Db 360 GGGGGCCCTCTGTTGAGTACCCCAAGCGAGAGCGGATGGTCTTAATGAAGACAGTAGA 419  
QY 1430 AGAAGAGGACCTAGAGATTGAGAGGCTTAAGACGAAGCAAAAAGAACTGGAGGCCAAGAT 1489  
Db 420 AGAAGAGGACCTAGAGATTGAGAGGCTTAAGACGAAGCAAAAAGAACTGGAGGCCAAGAT 479  
QY 1490 GTTGGCCCAAGAA 1501  
Db 480 GTTGGCCCAAGAA 491

RESULT 13  
AAS38879  
ID AAS38879 standard; cDNA; 386 BP.  
XX AAS38879;  
AC AAS38879;  
XX 17-DEC-2001 (first entry)  
DT  
XX Novel human diagnostic and therapeutic gene #1937.  
DE  
XX Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX PN WO200166753-A2.  
XX PD 13-SEP-2001.  
XX  
XX 09-MAR-2001; 2001WO-US007787.  
XX PF  
XX 09-MAR-2000; 2000US-0188609P.  
XX PR  
XX (CHIR ) CHIRON CORP.  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;  
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
XX Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;  
DR WPI; 2001-530177/58.  
XX  
XX New polynucleotides and polypeptides, useful for diagnosis and treatment  
PT of breast, lung and colon cancer.  
XX  
XX Claim 1; Page 1080; 1193pp; English.  
XX  
XX The invention relates to new polynucleotides and polypeptides, useful for  
CC diagnosis and treatment of breast, lung and colon cancer. The sequences  
CC can be used in detecting differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample derived from a  
CC cell suspected of being cancerous. They can also be used to inhibit  
CC tumour growth by modulating expression of a gene product. AAS36943-  
CC AAS39338 represent novel human diagnostic and therapeutic coding  
CC sequences of the invention  
XX  
XX Sequence 386 BP; 105 A; 102 C; 118 G; 61 T; 0 U; 0 Other;  
SQ

Query Match 24.5%; Score 376.4; DB 4; Length 386;  
Best Local Similarity 98.4%; Pred. No. 3e-93;  
Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1147 CTTGGGACCTTAAAGCTGCTCAGAAAGATTGCTGGTCCACAGAGCAAGAGAG 1206  
Db 1 CTTGGGACCTTAAAGCTGCTCAGAAAGATTGCTGGTCCACAGAGCAAGAGAG 60  
QY 1207 CCCGAGGCCCTGAGGAGGAGATTGGAGCCCTGAGCCATGCGAGCTCCAGCTCTG 1266  
Db 61 CCCGAGGCCCTGAGGAGGAGATTGGAGCCCTGAGCCATGCGAGCTCCAGCTCTG 120  
QY 1267 CTTCCAGAACTCAGCCCTTACAGAGCTTAAGCAGCATGACCCGCCATGCTGGAGC 1326  
Db 121 CTTCCAGAACTCAGCCCTTACAGAGCTTAAGCAGCATGACCCGCCATGCTGGAGC 180  
QY 1327 GCCTCTCAGCTTGAGCGCTGCTGCTCCAGGGAGCCAGGGGCCCTCTGTGA 1386  
Db 181 GCCTCTCAGCTTGAGCGCTGCTGCTCCAGGGAGCCAGGGGCCCTCTGTGA 240  
QY 1387 GTACCCCAAGCGAGCGGATGGTGTGTAATGAACAGACAGTGAAGAGACCTAGAGA 1446  
Db 241 GTACCCCAAGCGAGCGGATGGTGTGTAATGAACAGACAGTGAAGAGACCTAGAGA 300  
QY 1447 TTGAGAGCTTAAGACGAGCAAGCAAAAGAACTGGAGGCCAAGATGTTGCCCAAGGCTG 1506  
Db 301 TTGAGAGCTTAAGACGAGCAAGCAAAAGAACTGGAGGCCAAGATGTTGCCCAAGGCTG 360  
QY 1507 AGGAAAGGAGACCATTTGCCACA 1532  
Db 361 AGGAAAGGAGACCATTTGCCACA 386

RESULT 14  
ACH44493  
ID ACH44493 standard; cDNA; 464 BP.  
XX  
XX ACH44493;  
XX  
XX 13-OCT-2003 (first entry)  
XX  
XX Human foetal brain cDNA #5218.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX

PD 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 31705; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030073623  
XX  
XX Sequence 464 BP; 96 A; 136 C; 131 G; 96 T; 0 U; 5 Other;  
SQ

Query Match 17.2%; Score 265; DB 9; Length 464;  
Best Local Similarity 77.6%; Pred. No. 1.7e-62;  
Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 200 GCGGGCATGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCT 259  
Db 37 GCCTGGAATTCATATCGCGTTCGTAGAGTGTAAACGCGCCGAAACCAACCGATATCT 96  
QY 260 CAAATACAGTTTGTATGCTTTATGGGAGAGAGTACTCAGCAGGACATCTATGAGG 319  
Db 97 CCCTTACCTGTGCTTGTGCTTTGACAGCATAGAGGACTGCGCTGACATGTACTCTCC 156  
QY 320 TTCAGTCAGCCCATCTTAAGGCATCTTGGAGGGCAGATGCCATGTGCTTGCCTA 379  
Db 157 TTGACCGTGGACCGTGGCCACTGCCCCTCTTTCGTAAATAAAGCGGGAGCTGGCCTA 216  
QY 380 TGGACCCACAGAGCTGGGAGAGCGCACATCTGCGGAGCCCGAGAGCAACCTGGGGT 439  
Db 217 CGGTGCTCTGAGAGCTGGCATTAC-ATCTCGATGCTTGGCAGCCCGAGAGCAACCTGGGGT 275  
QY 440 GATCCCGGGGCTCTCATGACCTTCCTGACCTCAAGAGGAGAGGTGCGGAGGGCG 499  
Db 276 GATCCCGGGGCTCTCATGACCTTCCTGACCTCAAGAGGAGAGGTGCGGAGGGCG 335  
QY 500 GCCATGGCCCTTCTGTCAACATGTCTTACCTAGAGATCTACCAGGAGAGGTATTAGA 559  
Db 336 GCCATGGCCCTTCTGTCAACATGTCTTACCTAGAGATCTACCAGGAGAGGTATTAGA 395  
QY 560 CCTCCTGGACCCCTGCTTTCGGGAGACCTGGTAAATCCGAGAGACTGCCCGGGGAATATCCT 619  
Db 560 CCTCCTGGACCCCTGCTTTCGGGAGACCTGGTAAATCCGAGAGACTGCCCGGGGAATATCCT 619

Db 396 CCTCTGGACCTCTTGGGAGACCTTGGATCCGAGAGACTGCGGGGAATATCCT 455  
QY 620 GATTCCGGG 628  
Db |||||  
456 GATTCCGGG 464

RESULT 15

ACH77408  
ID ACH77408 standard; DNA; 531 BP.  
XX AC ACH77408;  
XX AC ACH77408;  
DT 29-JUL-2004 (first entry)  
XX Human genome derived single exon probe #10603.  
DE Human; probe; ss; gene expression; single exon probe; microarray;  
XX alternative splicing event; genomic alteration.  
XX Homo sapiens.  
OS  
XX US2003194704-A1.  
XX 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
XX (RANK/) RANK D R.  
XX (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX  
XX New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.

XX Claim 15; SEQ ID NO 10603; 80pp; English.  
XX  
XX The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 688 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross

CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX SQ Sequence 531 BP; 117 A; 153 C; 149 G; 112 T; 0 U; 0 Other;

Query Match 15.2%; Score 233.2; DB 12; Length 531;  
Best Local Similarity 98.7%; Pred. No. 1e-53;  
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 758 CAAGGTGGACCGGGAACGTTTGGCCCCCATTTCCGACGAGAGGAAAACCTTACCT 817  
Db 34 CCAGGTGGACCGGGAACGTTTGGCCCCCATTTCCGACGAGAGGAAAACCTTACCT 93  
QY 818 GATTGACTTGGCTGGGTTCAGAGGACAACCGGGGCACAGGCAACAAGGGCTTCGGCTAAA 877  
Db 94 GATTGACTTGGCTGGGTTCAGAGGACAACCGGGGCACAGGCAACAAGGGCTTCGGCTAAA 153  
QY 878 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATCGCTGAA 937  
Db 154 AGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATCGCTGAA 213  
QY 938 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCTCGCTTATTCGAGGAC 995  
Db 214 TCAGGGCCCTCCCTCGTGTACCTTATCGGACAGCAAGCTCAGCTCGCTTATTCGAGGAC 271

Search completed: November 10, 2004, 06:57:53  
Job time : 798.24 secs



Db 361 ATCCAGTGTGCTTGTCTTATGACCCACAGAGCTGGGAAGCGCACAAATGCTGGGCA 420  
QY 421 GCCAGAGCAACCTGGGGTATCCCGGGGCTCTCATGAGACCTCTGACGCTCACAGGG 480  
Db 421 GCCAGAGCAACCTGGGGTATCCCGGGGCTCTCATGAGACCTCTGACGCTCACAGGG 480  
QY 481 AGGAGGTGCGAGGGGCGGCCATGGGCCCTTTCTGTACCATGTCTTACTAGATCT 540  
Db 481 AGGAGGTGCGAGGGGCGGCCATGGGCCCTTTCTGTACCATGTCTTACTAGATCT 540  
QY 541 ACCAGAGAAAGTATTAGACCTCTCTGAGACCTCTCTCGGAGACCTGGTAAATCCGAGAAG 600  
Db 541 ACCAGAGAAAGTATTAGACCTCTCTGAGACCTCTCTCGGAGACCTGGTAAATCCGAGAAG 600  
QY 601 ACTGCCGGGGGAATATCTGATTCGGGTCTCTCCCAAGAGCCCATCAGTAGCTTTGCTG 660  
Db 601 ACTGCCGGGGGAATATCTGATTCGGGTCTCTCCCAAGAGCCCATCAGTAGCTTTGCTG 660  
QY 661 ATTTTGAGCGGCACTTCTCGGAGCCAGTCCGAATCGGACTGTAGAGCCACCCGGCTCA 720  
Db 661 ATTTTGAGCGGCACTTCTCGGAGCCAGTCCGAATCGGACTGTAGAGCCACCCGGCTCA 720  
QY 721 ACCAGGCTCTCTCCGCGATCATGCTGTCTCTGCTCAAGGTGGACCGGGAAGCTT 780  
Db 721 ACCAGGCTCTCTCCGCGATCATGCTGTCTCTGCTCAAGGTGGACCGGGAAGCTT 780  
QY 781 TGGCCCATTTCCGACAGAGAGGGAATACTTACCTGATTGCTTGGTGGGTCTAGAG 840  
Db 781 TGGCCCATTTCCGACAGAGAGGGAATACTTACCTGATTGCTTGGTGGGTCTAGAG 840  
QY 841 ACAACGGGCGCACAGGCAACAGGGCTTCGGCTTAAAGAGTGGAGCCATCAACCT 900  
Db 841 ACAACGGGCGCACAGGCAACAGGGCTTCGGCTTAAAGAGTGGAGCCATCAACCT 900  
QY 901 CCCTGTTTGTCTGGGCAAGTGGTAGATCGCTGATCAGGGCTCCCTCGGTACCTT 960  
Db 901 CCCTGTTTGTCTGGGCAAGTGGTAGATCGCTGATCAGGGCTCCCTCGGTACCTT 960  
QY 961 ATCGGAGACAGACTCACTCGCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTA 1020  
Db 961 ATCGGAGACAGACTCACTCGCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTA 1020  
QY 1021 TCCTTATGCGCAATGCCCCCTGAGAGCGCTTCTACCTAGACAGTCTCCGCACTCA 1080  
Db 1021 TCCTTATGCGCAATGCCCCCTGAGAGCGCTTCTACCTAGACAGTCTCCGCACTCA 1080  
QY 1081 ACTTTGTGCGCAGGTCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAG 1140  
Db 1081 ACTTTGTGCGCAGGTCCAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTCGAG 1140  
QY 1141 CTATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGTATGCTTGGTTCACACAGAGCAA 1200  
Db 1141 CTATGCTTTGGGACCTGTTAAGCTGTCTCAGAAAGTATGCTTGGTTCACACAGAGCAA 1200  
QY 1201 AGAGAGCCGAGCCCTGAGGAGAGGAGATTGGGAGCCCTGAGCCATGGAGCTCCAG 1260  
Db 1201 AGAGAGCCGAGCCCTGAGGAGAGGAGATTGGGAGCCCTGAGCCATGGAGCTCCAG 1260  
QY 1261 CCTCTGCTCCCAAGAACTCAGGCCCTTACAGAAAGCTTAAGCAGCATGGACCCCGCCATGC 1320  
Db 1261 CCTCTGCTCCCAAGAACTCAGGCCCTTACAGAAAGCTTAAGCAGCATGGACCCCGCCATGC 1320  
QY 1321 TGGAGGCTCTCTCAGTTGGAGCGTGTGCTTCCAGGGGAGCAGGGGGCCCTC 1380  
Db 1321 TGGAGGCTCTCTCAGTTGGAGCGTGTGCTTCCCTCCAGGGGAGCAGGGGGCCCTC 1380  
QY 1381 TGTGTAGTACCCAAAGCAGAGAGCGATGTGTCTAATGAAGACAGTAGAAGAGAGACC 1440  
Db 1381 TGTGTAGTACCCAAAGCAGAGAGCGATGTGTCTAATGAAGACAGTAGAAGAGAGACC 1440  
QY 1441 TAGAGATTGAGGCTTTAAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCGA 1500

Db 1441 TAGAGATTGAGAGGCTTTAAGACGAAGCAAAAAGAACTGGAGGCCAAGATGTTGGCCCGA 1500  
QY 1501 AGCTGAGGAAAAGAGAACCAATTGTCACCAATGTGA 1538  
Db 1501 AGCTGAGGAAAAGAGAACCAATTGTCACCAATGTGA 1538  
RESULT 2  
US-10-093-317-3  
; Sequence 3, Application US/10093317  
; Patent No. 6762043  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
; FILE OF INVENTION: their use  
; FILE REFERENCE: 1044  
; CURRENT APPLICATION NUMBER: US/10/093,317  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 09/724,224  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1538  
; TYPE: DNA  
; ORGANISM: Human  
US-10-093-317-3  
Query Match 100.0%; Score 1538; DB 4; Length 1538;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCAGCCGCGGCGCTCGACGCGAGAGGCGAGCGGAGATGGCGGAGCTTACGCG 60  
Db 1 ATGCAGCCGCGGCGCTCGACGCGAGAGGCGAGCGGAGATGGCGGAGCTTACGCG 60  
QY 61 CGCGATCTCAGAGCTGCTCGCTGCTGCTAAGCAAGATTGAGTCTCTAGAGATTGCTTCCAC 120  
Db 61 CGCGATCTCAGAGCTGCTGCTGCTGCTAAGCAAGATTGAGTCTCTAGAGATTGCTTCCAC 120  
QY 121 CTCAGCTCGGTAAAGGTGGCTGCTGCTGCTGCGGCACTTTGTTGATGGAACAGCGGAG 180  
Db 121 CTCAGCTCGGTAAAGGTGGCTGCTGCTGCTGCGGCACTTTGTTGATGGAACAGCGGAG 180  
QY 181 CAAGTATCCCTCTGTGCGGGGATGAGAGCTGCTCTAGAGATTGCTTAACTGGA 240  
Db 181 CAAGTATCCCTCTGTGCGGGGATGAGAGCTGCTCTAGAGATTGCTTAACTGGA 240  
QY 241 GGAACCAACAGAGACTCTCAATACAGTTTGTGATGCTTATGGGAGAGGACTCTC 300  
Db 241 GGAACCAACAGAGACTCTCAATACAGTTTGTGATGCTTATGGGAGAGGACTCTC 300  
QY 301 AGCAGGACATCTATGAGGTTTCAAGTTCAGTCCAGCCCACTTCTGTAAGGCACTTCTGTAAGGGG 360  
Db 301 AGCAGGACATCTATGAGGTTTCAAGTTCAGTCCAGCCCACTTCTGTAAGGCACTTCTGTAAGGGG 360  
QY 361 ATGCCAGTGTGCTTATGAGCCACAGAGCTGGGAGAGCGCACAAATGCTGGGCA 420  
Db 361 ATGCCAGTGTGCTTATGAGCCACAGAGCTGGGAGAGCGCACAAATGCTGGGCA 420  
QY 421 GCCAGAGCAACTGGGGTGTATCCCGGGGCTCTCATGAGCTCTTACCAAGGG 480  
Db 421 GCCAGAGCAACTGGGGTGTATCCCGGGGCTCTCATGAGCTCTTACCAAGGG 480  
QY 481 AGGAGGTGCGGAGGCGCCCATGGGCTTCTGTCTCATGAGCTTCTTACCTAGAGATCT 540  
Db 481 AGGAGGTGCGGAGGCGCCCATGGGCTTCTGTCTCATGAGCTTCTTACCTAGAGATCT 540  
QY 541 ACCAGAGAGGTTATTAGACCTCTGAGCCCTGCTTCCGAGAGCTTGTAAATCCGAGAAG 600  
Db 541 ACCAGAGAGGTTATTAGACCTCTGAGCCCTGCTTCCGAGAGCTTGTAAATCCGAGAAG 600  
QY 601 ACTGCCGGGGGAATATCTGATTCGGGTCTCTCCCAAGAGCCCATCAGTAGCTTTGCTG 660



```

Db 601 ACTGCGGGGGAATATCTCTGATTCGGGGTCTCTCCAGAAAGCCATCAGTAGCTTTGGTG 660
Qy 661 ATTTTGGAGGGCACTTTCTGCGAGCAGTCGAAATCGGACTCTAGGAGCCACCGGCTCA 720
Db 661 ATTTTGGAGGGCACTTTCTGCGAGCAGTCGAAATCGGACTCTAGGAGCCACCGGCTCA 720
Qy 721 ACCAGCGCTCTCTCCGCGAGTCACTGTGTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTT 780
Db 721 ACCAGCGCTCTCTCCGCGAGTCACTGTGTGCTCTCTGCTCAAGGTGGACAGCGGGAACGTT 780
Qy 781 TGGCCCCATTTGCGCAGCAGAGGAGAACTCTACTGATTAAGTCTGCTGGTGGGTGAGAG 840
Db 781 TGGCCCCATTTGCGCAGCAGAGGAGAACTCTACTGATTAAGTCTGCTGGTGGGTGAGAG 840
Qy 841 ACAACCGCGCAGCAGCAACAGGCGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACT 900
Db 841 ACAACCGCGCAGCAGCAACAGGCGCTTTCGGCTAAAGAGAGTGGAGCCATCAACACT 900
Qy 901 CCTGTTTGTCTGCGGCAAGTGTGTAGTGCCTGAATCAGGCGCTCTCTGCTGTACCTT 960
Db 901 CCTGTTTGTCTGCGGCAAGTGTGTAGTGCCTGAATCAGGCGCTCTCTGCTGTACCTT 960
Qy 961 ATCGGACAGCAAGTCTACTGCTTATTCAGGACTCTCTGCTGGTGGTTCAGGCCACAGTA 1020
Db 961 ATCGGACAGCAAGTCTACTGCTTATTCAGGACTCTCTGCTGGTGGTTCAGGCCACAGTA 1020
Qy 1021 TCCTTATTTGCCAAATTCGCCCTGAGAGAGCGTCTTACTGACACAGCTCTCCGCACTCA 1080
Db 1021 TCCTTATTTGCCAAATTCGCCCTGAGAGAGCGTCTTACTGACACAGCTCTCCGCACTCA 1080
Qy 1081 ACTTTGCTGCCAGGCTTCAAGAGGTGATCAATTCGGGCTTTTACCAATGAGAGCTTGACG 1140
Db 1081 ACTTTGCTGCCAGGCTTCAAGAGGTGATCAATTCGGGCTTTTACCAATGAGAGCTTGACG 1140
Qy 1141 CTATGCTTGGGACCTGTTAAGTGTCTCAGAAAAGATTTGCTGGTCCACAGAGGCAA 1200
Db 1141 CTATGCTTGGGACCTGTTAAGTGTCTCAGAAAAGATTTGCTGGTCCACAGAGGCAA 1200
Qy 1201 AGAGAGCGGAGGCGCTGAGGAGAGGAGATTTGGAGCGCTGAGCGCCATGCGAGTCCAG 1260
Db 1201 AGAGAGCGGAGGCGCTGAGGAGAGGAGATTTGGAGCGCTGAGCGCCATGCGAGTCCAG 1260
Qy 1261 CCTGCTCTCCAGAAACTCAGACCCCTTACAGAGCTAAGCAGCATGACCCCGGCAATGC 1320
Db 1261 CCTGCTCTCCAGAAACTCAGACCCCTTACAGAGCTAAGCAGCATGACCCCGGCAATGC 1320
Qy 1321 TGGAGCGCTCTCTAGCTTTGAGACGCTGCTTGGTCTCCAGGGAGCGAGGGGGGCGCTC 1380
Db 1321 TGGAGCGCTCTCTAGCTTTGAGACGCTGCTTGGTCTCCAGGGAGCGAGGGGGGCGCTC 1380
Qy 1381 TGTGAGTACCCCAAGCGAGGCGATGGTGTCTAATGACAGCATAGAGAGGAGGACC 1440
Db 1381 TGTGAGTACCCCAAGCGAGGCGATGGTGTCTAATGACAGCATAGAGAGGAGGACC 1440
Qy 1441 TAGAGATTGAGAGGCTTAAAGCAGAAAGAACTGGAGCCCAAGATTTGGGCCAGA 1500
Db 1441 TAGAGATTGAGAGGCTTAAAGCAGAAAGAACTGGAGCCCAAGATTTGGGCCAGA 1500
Qy 1501 AGGCTGAGGAAAGAGAACCAATTGTCACCAATGTGA 1538
Db 1501 AGGCTGAGGAAAGAGAACCAATTGTCACCAATGTGA 1538

```

RESULT 3

```

US-09-595-684B-34
; Sequence 34, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni

```

```

; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cycop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Human
; US-09-595-684B-34

```

```

Query Match 99.5%; Score 1530; DB 4; Length 2097;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 6 GCCGGGGGGGCTCGACCGCAGCAGAGGCGACGCGAGATGGCGGAGCTTCAGCGGGCGG 65
Db 28 GCCGGGGGGGCTCGACCGCAGCAGAGGCGACGCGAGATGGCGGAGCTTCAGCGGGCGG 87
Qy 66 ATCTCAGGAGCTGGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCTGTCGTCACCTCCA 125
Db 88 ATCTCAGGAGCTGGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCTGTCGTCACCTCCA 147
Qy 126 GCTCGCTAAGGGTGGCTGTGCGACTGCGGCCATTTTGGATGGAAACAGCGGAGCAAGT 185
Db 148 GCTCGCTAAGGGTGGCTGTGCGACTGCGGCCATTTTGGATGGAAACAGCGGAGCAAGT 207
Qy 186 GATCCCCCTGTGCGGGGATGAGAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAAAC 245
Db 208 GATCCCCCTGTGCGGGGATGAGAGCTGCTCTTAGAGATTGCTTAACCTGGAGGAAAC 267
Qy 246 CACGAGAGAGCTCTCAAAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 305
Db 268 CACGAGAGAGCTCTCAAAATACAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 327
Qy 306 GACATCTATGAGGTTTCACTGAGCCCAATCTTAAGGCACTTGTGTAAGGGGAGGAGTGC 365
Db 328 GACATCTATGAGGTTTCACTGAGCCCAATCTTAAGGCACTTGTGTAAGGGGAGGAGTGC 387
Qy 366 AGTGTGCTTGGCTATGAGCCCAAGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCCA 425
Db 388 AGTGTGCTTGGCTATGAGCCCAAGAGCTGGGAGAGCGCACAAATGCTGGGAGGCCCA 447
Qy 426 GAGCAACTGGGGTGTATCCCGGGCTCTCATGAGCTTCTGAGCTCACAAGGGAGGAG 485
Db 448 GAGCAACTGGGGTGTATCCCGGGCTCTCATGAGCTTCTGAGCTCACAAGGGAGGAG 507
Qy 486 GGTCCGAGGGCGGCCCATGCGGCCCTTTTGTGACCAATGCTTACCTTAGAGATCTACCAG 545
Db 508 GGTCCGAGGGCGGCCCATGCGGCCCTTTTGTGACCAATGCTTACCTTAGAGATCTACCAG 567
Qy 546 GAGAAGGTATTAGACCTCTCGGAGCCCTGCTTGGGAGAGCTGTTAATCCGAGAAAGACTGC 605
Db 568 GAGAAGGTATTAGACCTCTCGGAGCCCTGCTTGGGAGAGCTGTTAATCCGAGAAAGACTGC 627
Qy 606 CGGGGGAATATCTGATTTCCGGGCTCTCCAGAGCCCATCAGTAGCTTTTGTCTGATTTT 665
Db 628 CGGGGGAATATCTGATTTCCGGGCTCTCCAGAGCCCATCAGTAGCTTTTGTCTGATTTT 687
Qy 666 GAGCGGCACTTCTGCGGAGCCAGTCTGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 725
Db 688 GAGCGGCACTTCTGCGGAGCCAGTCTGAAATCGGACTGTAGGAGCCACCGGCTCAACCCAG 747
Qy 726 CGCTCTCTCCGAGCTCATGCTGCTCTCTGCTCAAGGTGGAGCCAGCGGGAAGCTTTGGCC 785
Db 748 CGCTCTCTCCGAGCTCATGCTGCTCTCTGCTCAAGGTGGAGCCAGCGGGAAGCTTTGGCC 807

```

Qy	786	CCATTTCCGACGAGAGGGAAAACTCTACCTGATTCACCTTGGCTGGGTGAGAGCAAC	845
Db	808	CCATTTCCGACGAGAGGGAAAACTCTACCTGATTCACCTTGGCTGGGTGAGAGCAAC	867
Qy	846	CGGCGCACGACACAGAGGCCCTTCGGCTAAAGAGAGATGGAGCCATCAACACCTCCCTG	905
Db	868	CGGCGCACAGCAACAAGGCCCTTCGGCTAAAGAGAGATGGAGCCATCAACACCTCCCTG	927
Qy	906	TTTGTCTCTGGGCAAAGTGGTAGATGCCTGAATCAGGGCCTCCCTCGTGTACCTTATTCGG	965
Db	928	TTTGTCTCTGGGCAAAGTGGTAGATGCCTGAATCAGGGCCTCCCTCGTGTACCTTATTCGG	987
Qy	966	GACAGCAAGCTCACTCCGCTATTGCAGGACTCTCTGGGTGGCTCAGCCCAAGTATCCCTT	1025
Db	988	GACAGCAAGCTCACTCCGCTATTGCAGGACTCTCTGGGTGGCTCAGCCCAAGTATCCCTT	1047
Qy	1026	ATTGCCAAACATTGGCCCTTGAGAGACGCTTCTACCTAGACACAGTCTCTCGCACTCAACTTT	1085
Db	1048	ATTGCCAAACATTGGCCCTTGAGAGACGCTTCTACCTAGACACAGTCTCTCGCACTCAACTTT	1107
Qy	1086	GCTGCCAGTCCAAAGAGGTGATCAATCGGCCTTTTACCAATGAGAGCCTGCAAGCCTCAT	1145
Db	1108	GCTGCCAGTCCAAAGAGGTGATCAATCGGCCTTTTACCAATGAGAGCCTGCAAGCCTCAT	1167
Qy	1146	GCCTTGGGACCTGTTTAAGCTGTCTCAAGAAAGAAATTTGCTTGGTCCACAGAGGCAAGAGA	1205
Db	1168	GCCTTGGGACCTGTTTAAGCTGTCTCAAGAAAGAAATTTGCTTGGTCCACAGAGGCAAGAGA	1227
Qy	1206	GCCGAGGCCCTGAGGAGAGGAGATTTGGAGCCCTGAGGCCATGGCAGCTCAGACCTCT	1265
Db	1228	GCCGAGGCCCTGAGGAGAGGAGATTTGGAGCCCTGAGGCCATGGCAGCTCAGACCTCT	1287
Qy	1266	GCCTCCCAAGAACTCAGCCCCCTACAGAACTAAGACATGGAACCCGGCCATGCTGGAG	1325
Db	1288	GCCTCCCAAGAACTCAGCCCCCTACAGAACTAAGACATGGAACCCGGCCATGCTGGAG	1347
Qy	1326	CGCCTCTCAGCTTGGACCGTCTGCTTTCCTCCAGGGGAGCCAGGGGGCCCTCTGTGTG	1385
Db	1348	CGCCTCTCAGCTTGGACCGTCTGCTTTCCTCCAGGGGAGCCAGGGGGCCCTCTGTGTG	1407
Qy	1386	AGTACCCCAAGCGAGCGGATGGTGCTTAATGAGNACGTAGNAGNAGAGGACCTTAGAG	1445
Db	1408	AGTACCCCAAGCGAGCGGATGGTGCTTAATGAGNACGTAGNAGNAGAGGACCTTAGAG	1467
Qy	1446	ATTGAGAGGCTTAAGACCAAGCAAAAAGAACTGGAGCCCAAGATGTTGGGCCCAAGAGGCT	1505
Db	1468	ATTGAGAGGCTTAAGACCAAGCAAAAAGAACTGGAGCCCAAGATGTTGGGCCCAAGAGGCT	1527
Qy	1506	GAGGAAAGAGAACCAATGTCACCAATG	1535
Db	1528	GAGGAAAGAGAACCAATGTCACCAATG	1557

## RESULT 4

```

US-09-724-224-7
; Sequence 7, Application US/09724224
; Patent No. 6387644
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6387644el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1044
; CURRENT APPLICATION NUMBER: US/09/724,224
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-09-724-224-7

```

Query Match		95.0%; Score 1461; DB 3; Length 1464;	
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	78	G G T C G C T G T C G G C T T A G C A A G A T T G G A G C T A C T C G T G C T C A C C T C C A G C T C C G G T A A G G	137
Db	4	G G T C G C T G T C G G C T A A G C A A G A T T G G A G C T A C T C G T G C T C A C C T C C A G C T C C G G T A A G G	63
Qy	138	G T G C C T G T G C G A C T G C G C C C A T T T G T G G A T G G A A C A G C G G G A G C A A G T G A T C C C C C C T G T	197
Db	64	G T G C T G T G C A C T G C G G C C A T T T G T G G A T G G A A C A G C G G G A G C A A G T G A T C C C C C C T G T	123
Qy	198	G T G C G G G C A T G A C A G C T G C T C T A G A G A T T G C T A A C T T G G A G G A A C C A C C A G G A G A C T	257
Db	124	G T G C G G G C A T G A C A G C T G C T C T A G A G A T T G C T A A C T T G S A G G A A C C A C C A G G A G A C T	183
Qy	258	C T C A A T A C C A G T T T G A T G C T T C T A T G G G A G A G G A G T A C T C A G C A G G A C A T C T A T G C A	317
Db	184	C T C A A T A C C A G T T T G A T G C T T C T A T G G G A G A G G A G T A C T C A G C A G G A C A T C T A T G C A	243
Qy	318	G G T T C A G T C A G C C C A P C C T A A G G C A C T T G C T G A A G G G C A G A A T G C C A G T G C T T G C C	377
Db	244	G G T T C A G T C A G C C C A P C C T A A G G C A C T T G C T G A A G G G C A G A A T G C C A G T G C T T G C C	303
Qy	378	T A T G G A C C C A C A G A G C T G G A A G A C G C A C A C A A T G C T G G G C A G C C C A G A G C A A C C T G G G	437
Db	304	T A T G G A C C C A C A G A G C T G G A A G A C G C A C A C A A T G C T G G G C A G C C C A G A G C A A C C T G G G	363
Qy	438	G T G A T C C C G C G G C T C T C A T G A G C A C C T C T G C A G C T C A A G A G G A G A G G T G C C G A G G C	497
Db	364	G T G A T C C C G C G G C T C T C A T G A G C A C C T C T G C A G C T C A A A G G A G A G G T G C C G A G G C	423
Qy	498	C G G C A T G G C C C T T T C T C A C A G T C T T A C C T A G A G A T C T A C C A G G A A G A G T A T T A	557
Db	424	C G G C A T G G C C C T T T C T G T C A C A G T C T T A C C T A G A G A T C T A C C A G G A A G A G G T A T T A	483
Qy	558	G A C C T C C T G A C C C T G C T T C G G A G A C C T G G T A A T C C G A A G A C T C C C G G G G G A A T A T C	617
Db	484	G A C C T C C T G A C C C T G C T T C G G A G A C C T G G T A A T C C G A A G A C T G C C G G G G G A A T A T C	543
Qy	618	C T G A T T C C G G G T C T C C C A A A G C C C A T C A G T A G C T T T G C T G A T T T G A G C G G C A C T T C	677
Db	544	C T G A T T C C G G G T C T C C C A A A G C C C A T C A G T A G C T T T G C T G A T T T G A G C G G C A C T T C	603
Qy	678	C T C C A G C A G T C G A A A T C G G A C T G T A G G A G C A C C C G C T C A A C A G A G C C T C C T C C C G C	737
Db	604	C T C C A G C A G T C G A A A T C G G A C T G T A G A G A C C A C C C G C T C A A C A G A G C T C C T C C C G C	663
Qy	738	A G T C A T C T G T G T C C T G G T C A A G G T G G A C C A G C G G G A A C G T T T G B C C C C A T T T C G C C A G	797
Db	664	A G T C A T C T G T G T C C T G G T C A A G G T G G A C C A G C G G G A A C G T T T G G C C C C A T T T C G C C A G	723
Qy	798	C G A G A G G A A A C T C T A C C T G A T T G A C T T G G C T G A G A G A C A A C C G G C C A C A G G C	857
Db	724	C G A G A G G A A A C T C T A C C T G A T T G A C T T G G C T G G G T C A G A G A C A A C C G G C C A C A G G C	783
Qy	858	A A C A A G G G C C T T C G G C T A A A A G A G A T G G A G C C A T C A A C A C C T C C C T G T T G T C C T G G G C	917
Db	784	A A C A A G G C C T T C G G C T A A A A G A G A T G G A G C C A T C A A C A C C T C C C T G T T G T C C T G G G C	843
Qy	918	A A A G T G T A G A T C G C T G A A T C A G G G C C T C C T C G T G T A C C T T A T C G G A C A G A C A G C T C	977
Db	844	A A A G T G T A G A T C G C T G A A T C A G G G C C T C C C T C G T G T A C C T T A T C G G A C A G A C A G C T C	903
Qy	978	A C T C G C C T A T T G C A G A C T C T C T G G T G G C T C A G C C C A C A G T A T C C T T A T T G C C A A C A T T	1037
Db	904	A C T C G C C T A T T G C A G A C T C T C T G G T G G C T C A G C C C A C A G T A T C C T T A T T G C C A C A T T	963
Qy	1038	G C C C C T G A G A G A C G C T T C T A C C T A G A C A C A G T C T C G C A C T C A A C T T T G C T C C A G G T C C	1097
Db	964	G C C C C T G A G A G A C G C T T C T A C C T A G A C A C A G T C T C G C A C T C A A C T T T G C T G C A G G T C C	1023

QY 1098 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGACGCTCATGCTTGGGACCT 1157  
 DB 1024 AAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGACGCTCATGCTTGGGACCT 1083  
 QY 1158 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT 1217  
 DB 1084 GTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGCAAGAGAGCCCGAGGCCCT 1143  
 QY 1218 GAGGAAGAGAGATTTGGGAGCCCTGAGCCCATGAGAGCTTCCAGCTTGCCTCCAGAAA 1277  
 DB 1144 GAGGAAGAGAGATTTGGGAGCCCTGAGCCCATGAGAGCTTCCAGCTTGCCTCCAGAAA 1203  
 QY 1278 CTCAGCCCTTACAGAACTAAGAGCATGAGAGCATGAGCCCGCCATGCTGGAGCCCTCTCAGC 1337  
 DB 1204 CTCAGCCCTTACAGAACTAAGAGCATGAGAGCATGAGCCCGCCATGCTGGAGCCCTCTCAGC 1263  
 QY 1338 TTGGACCGTCTGCTTGGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397  
 DB 1264 TTGGACCGTCTGCTTGGCTCCAGGGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323  
 QY 1398 CGAGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT 1457  
 DB 1324 CGAGAGCGGATGCTGCTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTT 1383  
 QY 1458 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGAGGCTGAGGAAAAGGAG 1517  
 DB 1384 AAGACGAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCAGAGAGGCTGAGGAAAAGGAG 1443  
 QY 1518 AACCAATTGTCACAAATGTGA 1538  
 DB 1444 AACCAATTGTCACAAATGTGA 1464

RESULT 5

US-10-093-317-7  
 ; Sequence 7, Application US/10093317  
 ; Patent No. 6762043  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: No. 6762043el motor proteins and methods for  
 ; TITLE OF INVENTION: their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/093,317  
 ; CURRENT FILING DATE: 2002-03-06  
 ; PRIOR APPLICATION NUMBER: 09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq For Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 1464  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-093-317-7

Query Match 95.0%; Score 1461; DB 4; Length 1464;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGTCTGCTGCTAAGCAAGATTGGAGTACTCTGCTGCACTCCAGCTCGCGTAAGG 137  
 DB 4 GGTCTGCTGCTAAGCAAGATTGGAGTACTCTGCTGCACTCCAGCTCGCGTAAGG 63  
 QY 138 GTGGCTGTGCGACTCGGCGCAATTTCTGGATGGAAACAGCGGGAGCAAGTGATCCCGCTGT 197  
 DB 64 GTGGCTGTGCGACTCGGCGCAATTTCTGGATGGAAACAGCGGGAGCAAGTGATCCCGCTGT 123  
 QY 198 GTGGCGGCGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 257  
 DB 124 GTGGCGGCGCATGGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 183  
 QY 258 CTCAAATACCAAGTTTGTGCTCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317  
 DB 184 CTCAAATACCAAGTTTGTGCTCTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243

QY 318 GGTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAAGGCGCAATGCGCAGTGTGCTTGGC 377  
 DB 244 GGTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAAGGCGCAATGCGCAGTGTGCTTGGC 303  
 QY 378 TATGAGCCCAAGAGAGCTGGGAAGACGCAACAATGCTGGGAGAGAGGAGCAACCTGGG 437  
 DB 304 TATGAGCCCAAGAGAGCTGGGAAGACGCAACAATGCTGGGAGAGAGGAGCAACCTGGG 363  
 QY 438 GTGATCCCGGGCTCTCATGAGACCTCTGAGCTCTCAAGGGAGAGGAGGAGGAGGAGG 497  
 DB 364 GTGATCCCGGGCTCTCATGAGACCTCTGAGCTCTCAAGGGAGAGGAGGAGGAGGAGG 423  
 QY 498 CGGCATCGGGCCCTTTCTGTGTCACCATGCTTACCTAGAGATCTACAGAGAGAGGATTA 557  
 DB 424 CGGCATCGGGCCCTTTCTGTGTCACCATGCTTACCTAGAGATCTACAGAGAGAGGATTA 483  
 QY 558 GACCTCTGAGACCTGCTTCCGAGAGACCTGTAATCCGAGAGAGAGCTGCGGGGGAATATC 617  
 DB 484 GACCTCTGAGACCTGCTTCCGAGAGACCTGTAATCCGAGAGAGAGCTGCGGGGGAATATC 543  
 QY 618 CTGATTCGGGCTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 677  
 DB 544 CTGATTCGGGCTCTCTCCAGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTC 603  
 QY 678 CTGCGCAGCAGTCGAATTCGAGCTGTAGAGAGCCACCGCTCAACAGAGGCTCTCCCGC 737  
 DB 604 CTGCGCAGCAGTCGAATTCGAGCTGTAGAGAGCCACCGCTCAACAGAGGCTCTCTCCCGC 663  
 QY 738 AGTCATGCTGCTCTGCTCAAGGTGCAAGCGGGAGACGTTTGGCCCCCAATTTGCCAG 797  
 DB 664 AGTCATGCTGCTCTGCTCAAGGTGCAAGCGGGAGACGTTTGGCCCCCAATTTGCCAG 723  
 QY 798 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTGAGGAGCAACCGGGGCGCAGGC 857  
 DB 724 CGAGAGGAGAACTCTACCTGATTTGACTTGGCTGAGGAGCAACCGGGGCGCAGGC 783  
 QY 858 AACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTCTGTTGCTCTGGGC 917  
 DB 784 AACAGGGCCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCTCTGTTGCTCTGGGC 843  
 QY 918 AAAGTGTGATGCTGCTGAATCAGGGCTCCTCTGCTGATACCTTATCGGGAGAGCAAGCTC 977  
 DB 844 AAAGTGTGATGCTGCTGAATCAGGGCTCCTCTGCTGATACCTTATCGGGAGAGCAAGCTC 903  
 QY 978 ACTGCTCTATTGCGAGGACTCTCTGGTGGCTCAGGCCACAGTATCTTATGCGCAACATT 1037  
 DB 904 ACTGCTCTATTGCGAGGACTCTCTGGTGGCTCAGGCCACAGTATCTTATGCGCAACATT 963  
 QY 1038 GCCCTGAGAGAGCGCTTCTACCTAGACACAGCTCTCCGCACTCAACTTTGCTGCGCAGGTCC 1097  
 DB 964 GCCCTGAGAGAGCGCTTCTACCTAGACACAGCTCTCCGCACTCAACTTTGCTGCGCAGGTCC 1023  
 QY 1098 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGAGCT 1157  
 DB 1024 AAGGAGGTGATCAATCGGCTTTTACCAATGAGAGCTTGCAGCTCATGCTTGGGAGCT 1083  
 QY 1158 GTTAAGCTGTCTCAGAAAGAAATTTGGTTCACAGAGGCAAGAGAGCCCGAGGCCCT 1217  
 DB 1084 GTTAAGCTGTCTCAGAAAGAAATTTGGTTCACAGAGGCAAGAGAGCCCGAGGCCCT 1143  
 QY 1218 GAGGAAGAGAGATTTGGGAGCCCTGAGCCCATGAGCTTCCAGCTCTGCTTCCAGAAA 1277  
 DB 1144 GAGGAAGAGAGATTTGGGAGCCCTGAGCCCATGAGCTTCCAGCTCTGCTTCCAGAAA 1203  
 QY 1278 CTCAGCCCTTACAGAACTAAGAGCATGAGAGCCCGGCTATGCTGGAGCGCTCTCTCAGC 1337  
 DB 1204 CTCAGCCCTTACAGAACTAAGAGCATGAGAGCCCGGCTATGCTGGAGCGCTCTCTCAGC 1263  
 QY 1338 TTGAGACCGTCTGCTTGCCTCCAGGGAGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1397  
 DB 1264 TTGAGACCGTCTGCTTGCCTCCAGGGAGGAGCCAGGGGGCCCTCTGTTGAGTACCCCAAG 1323







```

; OTHER INFORMATION: protein gene Hskip3a (Figure 1).
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a gene.
US-09-883-096-1

Query Match          9.0%; Score 138.6; DB 4; Length 4108;
Best Local Similarity 51.0%; Pred. No. 1.8e-29;
Matches 448; Conservative 0; Mismatches 404; Indels 27; Gaps 4;

Qy 270 TTTGATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAAGTTTCAGTGCAG 329
Db 344 TTTGACCGGCTTTTGGCAGGCGCCACCAACAGAGAGGTTTCCAGCACACACGAC 403

Qy 330 CCCATCTTAAGCACTTCTGGAAGGAGAGTCCAGTGTCTGCTTGTCTATGAGCCACA 389
Db 404 AGCGTCTGGACAGCTTCTCAGGGCTACAACTGCTCAGTGTGCTTGTGCTTGTGCTG 463

Qy 390 GGAGCTGGAGAGCAGACACATGCTGGCAGCCAGAGCACTTGGGGTATCCCGGG 449
Db 464 GGGGCTGGAGAGACACACACATGCTGGGAAGGAGGGGACCCCGGCATCATGTACCTG 523

Qy 450 GCTCTCATGACCTCTCTGAGCTCAAGGAGGAGGAGTGGCGAGGGCGGCCATGGGCC 509
Db 524 AC-----CACGCTGMACTGTACAGGCGCTGGAGGCGCCGACAGAGAGAGAC 574

Qy 510 CTTTCTGTCAACATGCTTACTAGATCTACAGAGAGAGTATTAAGACTCTCTGGAC 569
Db 575 TTCGAGGTGCTCATCAGCTACAGAGGTGTATATGAACAGATCCATGACCTCTCTGGAG 634

Qy 570 CTTGCTTGGGAGACCTGTTATCCGAGAGACTGCGGGGGATATCTGATTCGGGT 629
Db 635 C---CCAAAGGGGCGCTTGGCCATCCGAGAGACCCCGACAAAGGGGGTGGTGTCAAGGA 691

Qy 630 CTCTCCCAAGAGCCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTCTCTGCAAGCAGT 689
Db 692 CTTTCTTCCACAGCAGCTCAGCCAGCAGCTGCTGGAGATCTACACAGGGGGAGAC 751

Qy 690 CGAAATCGGACTGTAGAGACCAACCGGCTCAACAGCGCTCTCTCCGAGTCTATGCTGTG 749
Db 752 GGTAAACGACGACGACACCCACTGATGCCAAAGCGGACTTCTCTCCGCTCCCATGCCATC 811

Qy 750 CTCCTGGTCAAGGTGACCGAGGAGAGCGTTGGGCCCATTTTGGCCAG-----CGAGAG 803
Db 812 TTCAGATCTTTTGTGAAGCAGCAGGACCGGGTTCCAGGACTGACCCAGGCTGTCCAGGTG 871

Qy 804 GGAATACTCTACCTGATGACTTGGCTGGTTCAGAGGACAAACCGCGCAGCAGGCAACAAG 863
Db 872 CCAAGATGAGCTGATTTGACTGCTGGCTCAGAGCGGGCATCCAGCACCCATGCGAG 931

Qy 864 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTGTCTGGGCAAGTG 923
Db 932 GGGGAGCGGCTGCGGAGGGGGCCAAACATCAACCGCTCTCTGCTGGGCTCATCAACGTC 991

Qy 924 GTAGATGCGCTG-----AATCAGGGCTCCCTCTGTGTACCTTATTCGGGACAGCAG 974
Db 992 CTCATGCTTTGGCGGATGCAAGGGCGCGAAGACCCCATGTGCTCCATCCCGGACAGCAA 1051

Qy 975 CTCACCTCCCTATTGACGAGCTCTCTGGTGGCTCAGCCCAAGTATCTTTATTTGCCAAC 1034
Db 1052 CTGACCCGCTGCTCAAGAGATCTCCCTCGGGGGAACCTCCGACAGTATGATGCTGCTG 1111

Qy 1035 ATTGCCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAG 1094
Db 1112 ATCAGCCCTCCAGCCTGACCTACGAGGACAGCTACAAACACCCCTCAATATATGCCGAC 1171

Qy 1095 TCCAAGGAGGTGATCAATCGGCTTTTACCAATGAGAC 1133
Db 1172 GCCAAGGAGATCAGGCTCTCTGCTGAAGAGCAATGTGACC 1210

```

RESULT 11  
 US-09-883-096-6  
 ; Sequence 6, Application US/09883096  
 ; Patent No. 6680369

```

; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Craven, Andrew
; APPLICANT: Yu, Ming
; APPLICANT: Sakowicz, Roman
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
; TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE
; FILE REFERENCE: 020552-001410US
; CURRENT APPLICATION NUMBER: US/09/883,096
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 09/594,655
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hskip3a
; OTHER INFORMATION: fragment
; OTHER INFORMATION: Nucleotide sequence of Hskip3a fragment used in
; OTHER INFORMATION: ATPase assay (Figure 5).
US-09-883-096-6

Query Match          8.9%; Score 136.6; DB 4; Length 1152;
Best Local Similarity 51.2%; Pred. No. 3.5e-29;
Matches 437; Conservative 0; Mismatches 389; Indels 27; Gaps 4;

Qy 270 TTTGATGCTTCTATGGGAGGAGTACTCAGCAGGACATCTATGCAAGTTTCAGTGCAG 329
Db 214 TTTGACCGGCTTTTGGCAGGCGCCACCAACAGAGAGTGTTCACAGCACACACGAC 273

Qy 330 CCCATCTTAAGCACTTCTGGAAGGAGGAGTGGCGAGCCAGCAACTCTGGGGTGTATCCCGGG 449
Db 274 AGCGTCTGGACAGCTTCTCTCAGGGCTTAACTGCTCAGTGTGCTTGTGCTTGTGCTG 333

Qy 390 GGAGCTGGAGAGCAGCACAACTCTGGGAGCCAGAGCAACTCTGGGGTGTATCCCGGG 449
Db 334 GGGGCTGGAGAGACACACACCATCTGGAAGGAGGAGGGGACCCCGGCATCATGTA---- 389

Qy 450 GCTCTCATGAGACTCTCTGAGCTCAAGGAGGAGGAGTGGCGAGGGCGGCCCATGGGCC 509
Db 390 -----CTGACACCGTGGAACTGTACAGGCGCTGGAGGCGCCGACAGAGAGAC 444

Qy 510 CTTTCTGTCAACATGCTTACCTAGAGATCTACAGGAGAGGATTTAGACTCTCTGAC 569
Db 445 TTCGAGTGTCTCATCAGCTACAGGAGTGTATTAATGAACAGATCCATGACCTCTCTGGAG 504

Qy 570 CTTCTTTCGGGAGACCTGTTAATCCGAGAGACTGCGGGGGAATATCTGATTCGGGT 629
Db 505 C---CCAAAGGGGCGCTTGGCATCCGAGAGACCCCGACCAAGGGGGTGGTGTGAAGA 561

Qy 630 CTCTCCAGAGGCCATCAGTAGCTTTGCTGATTTTGAGCGGCACTTCTCTCCAGCCAGT 689
Db 562 CTTTCTTTCACAGCAGCCCTCAGCCAGCAGCTGTCTGGAGATCTTACCCAGGGGAA 621

Qy 690 CGAAATCGGACTGTAGAGAGCCACCGGCTCAACAGCGCTCTCTCCGAGTATGCTGTG 749
Db 622 CGTAAACGCGACGCGAGCACCACCTGATGCCAAACCGGACTTCTCTCCGCTCCCATGCCATC 681

Qy 750 CTCCTGTGCAAGGTGGACAGCGGGAAGCTTTGGCCCCCATTTTCCGCGAG-----CGAGAG 803
Db 682 TTCAGATCTTTTGAAGCAGCAGAGCCGGGTTCCAGGACTGACCCAGGCTGTCCAGGTG 741

Qy 804 GGAATACTCTACCTGATTTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCAGAGCAACAAG 863
Db 742 GCCAAGATGAGCTGTATTGACTGCTGGTGGCTCAGAGCGGGCATCCAGCACCCATGCGAAG 801

Qy 864 GGCCTTCGGCTAAAGAGAGTGGAGCCATCAACACTCCCTCTTGTCTGGGCAAGTG 923
Db 802 GGGGAGCGGCTGCGGAGGGGGCCAAACATCAACCGCTCTCTGCTGGCGCTCATCAACGTC 861

```







```

Db 226 CCCACAGGCTGGGAAACCTTACACCATGCTGGGCACAGACCAGGAGCTGGCATCTAT 285
Qy 444 CCGCGGGCTCTCATGACCTCTGAGCTCAACAGGGAGGAGGTCGCGAGGGCGGCA 503
Db 286 GTTCAGACCTCAACGACCTCTTCGTCGCATCGAGGAGACAGCAATGA -----C 336
Qy 504 TGGGCCCCCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTC 563
Db 337 ATGAGATGATGAGTCTCCATGCTCTACCTGAGATCTACATGATGATATCGGGACCTG 396
Qy 564 CTGGACCTCTCTGGGAGACCTGTAATCCGAGAGATCTACAGGAGAGGTATTAGACCTC 623
Db 397 CTGAACCCCTCCCTGGGCTACCTGGAGCTCGGGAGGACTCTAAAGGGGGTATCCAGGTG 456
Qy 624 CCGGCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTCTCTGCCA 683
Db 457 GCGGCTATCCGAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Qy 684 GCCAGTCAAAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCTCTCCCGCAGTCTAT 743
Db 517 GGGAAACCGGAGAGGACCCAGAGCCACGCGCCCAACAGAGCTCTCCCGCTCCAC 576
Qy 744 GCTGTCTCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTTCGCCAGGAG 803
Db 577 GCGGCGCCCTGTTCA ----- 592
Qy 804 GGAATACTCTACCTGATTTGCTGGTTCAGAGAACCCGCGCAGAGGCAAG 863
Db 593 -----TGATCGACTGGCTGGCTCAGAGCGCGCTCGCAGACACAGAAATCGT 639
Qy 864 GGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGCAAGTG 923
Db 640 GGGCAGCTATGAAGGAGGGGCCACATCAACGCTCAGTCTGTCGCTGGGCAATGCG 699
Qy 924 GTAGATCGCTGATATCAG-----GGCCTCCCTGCTGTACCTTATCGGGACAGCAAGCTC 977
Db 700 ATCAACGCCCTTGACGACAGGGTAGCAACAAGTACATCACTATCGCAGCAGAGCTC 759
Qy 978 ACTGCTCTATGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCCAATTT 1037
Db 760 ACCGGCTCTGAGGACTCTCTGGGAGAAACAGCCGACAGTATGATCGTCAATC 819
Qy 1038 GCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTC 1097
Db 820 AGTCTCGAGAGTGTCTTCGAGAGTCCCGGAACACCTTGACCTACGCGCGCGGCC 879
Qy 1098 AAGAGGTGATCAATCGG 1115
Db 880 AAGAACAATTAAGACTAGG 897

```

RESULT 14

```

US-09-724-516-1
; Sequence 1, Application US/09724516
; Patent No. 6391573
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6391573el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1045
; CURRENT APPLICATION NUMBER: US/09/724,516
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/621,233
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Human
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (472)...(495)
; OTHER INFORMATION: n = a, c, t, or g
US-09-724-516-1

Query Match      8.4%; Score 128.8; DB 3; Length 897;
Best Local Similarity 50.1%; Pred. No. 5.4e-27;
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;

Qy 264 TACCAGTTTGTATGCTTCTATGGGAGAGGAGTACTACAGCAGGACATCTATGACGTTCA 323
Db 106 TACCTGTTTCGACGTGGGCTTTGACTTACCCGCCAGGAGATGGTGTATCAGGCCACC 165
Qy 324 GTGACGCCCATCTTAAGGCACTTGTGAAGGAGCAATGCGAGTGTCTTGCCTTATGA 383
Db 166 ACCAAGACCTCATCGAGGGCGTCTATCTCAGGCTACATGCACTGTCTTTGCCATGGC 225
Qy 384 CCCACAGAGCTGGGAAGACGACACAATGCTGGGCAGCCAGAGCAACCTGGGGTGATC 443
Db 226 CCCACAGGCTGTGGGAAACCTACACCATGCTGGGCACAGACAGGAGCTTGGCATCTAT 285
Qy 444 CCGGGGCTCTCATGAGACCTCTCTGAGCTCAAGGAGAGGGTGGCCGAGGGCCGGCA 503
Db 286 GTTCAGACCCCTCAACGACCTCTTCCGTGCCATCGAGGAGACCAATGA -----C 336
Qy 504 TGGGCCCTTTCTGTCAACCATGCTTACCTAGAGATCTACAGGAGAGGTATTAGACCTC 563
Db 337 ATGAGATGATGAGTCTCCATGCTCTACCTGAGATCTACATGATGATATCGGGACCTG 396
Qy 564 CTGACCTCTGCTCGGAGACCTGTATATCCGAGAAGCTCCCGGGGAATATCTCTGATT 623
Db 397 CTGAACCCCTCCCTGGGCTACTGAGCTGGGAGGACTCTAAGGGGGTGTATCCAGGTG 456
Qy 624 CCGGCTCTCTCCAGAACCCATCAGTAGCTTTGCTGATTTTGGCGGCACCTCTCTGCCA 683
Db 457 GCGGCTATCCGAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 516
Qy 684 GCCAGTCAAAATCGGACTGTAGGAGCCACCGGCTCAACAGAGCTCTCCCGCAGTCTAT 743
Db 517 GGGAAACCGGAGAGGACCCAGGAGCCACGCGCCCAACAGAGCTCTCCCGTCCAC 576
Qy 744 GCTGTCTCTCTGCTCAAGGTGACAGCGGGAACGTTTGGCCCCATTTTCGCCAGGAG 803
Db 577 GCGGCGCCCTGTTCA ----- 592
Qy 804 GGAATACTCTACCTGATTTGCTGGTTCAGAGAACCCGCGCAGAGGCAAG 863
Db 593 -----TGATCGACTGGCTGGCTCAGAGCGCGCTCGCAGACACAGAAATCGT 639
Qy 864 GGCCTTCGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGCAAGTG 923
Db 640 GGGCAGCTATGAAGGAGGGGCCACATCAACGCTCAGTCTGTCGCTGGGCAATGCG 699
Qy 924 GTAGATCGCTGATATCAG-----GGCCTCCCTGCTGTACCTTATCGGGACAGCAAGCTC 977
Db 700 ATCAACGCCCTTGACGACAGGGTAGCAACAAGTACATCACTATCGCAGCAGAGCTC 759
Qy 978 ACTGCTCTATGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATTTGCCAATTT 1037
Db 760 ACCGGCTCTGAGGACTCTCTGGGAGAAACAGCCGACAGTATGATCGTCAATC 819
Qy 1038 GCCCTGAGAGCGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGGTC 1097
Db 820 AGTCTCGAGAGTGTCTTCGAGAGTCCCGGAACACCTTGACCTACGCGCGCGGCC 879
Qy 1098 AAGAGGTGATCAATCGG 1115
Db 880 AAGAACAATTAAGACTAGG 897

```

RESULT 15

```

US-10-090-695-1
; Sequence 1, Application US/10090695

```

; Patent No. 6664072  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: No. 6664072el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1045  
; CURRENT APPLICATION NUMBER: US/10/090,695  
; CURRENT FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: 09/621,233  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 897  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (472)...(495)  
; OTHER INFORMATION: n = a, c, t, or g  
US-10-090-695-1

Query Match 8.4%; Score 128.8; DB 4; Length 897;  
Best Local Similarity 50.1%; Pred. No. 5.4e-27;  
Matches 430; Conservative 0; Mismatches 356; Indels 72; Gaps 3;  
QY 264 TACAGTTTGCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAAGTTCA 323  
DB 106 TACCTGTTCAGCTGGGCTTTGACTTCACCGCCACCAGGAGATGTTATCAGGCCACC 165  
QY 324 GTGCAGCCCATCTTAAGGCACTTGTGGAAGGCGAGAAATGCCAGTGTGCTTGCCTATGGA 363  
DB 166 ACCAAGAGCTCATCGAGGCGTCACTCAGGCTACATGCCACTGTCTTTGGCTATGGC 225  
QY 384 CCACAGGAGCTGGGAAGACGACACAAATCTGGGAGCCAGAGCAACCTGGGGTGATC 443  
DB 226 CCACAGGCTGTGGGAAACCTACACCATCTGGGACAGCAGGAGCCTGGCACTAT 285  
QY 444 CCGGGGCTCTATGAGACCTCTGAGCTCAAGGAGGAGGAGGTCGCGAGGGCCGCCA 503  
DB 286 GTTCAGACCCCTAACGACCTCTTCCGTGCCATCGAGGAGACCAATGA-----C 336  
QY 504 TGGGCCCTTCTGTACCATGTCTTACCTAGAGATCTACCAGGAGAGGTATTAGACCTC 563  
DB 337 ATGGAGTATGAGTCTTCATGTCTTACCTGAGATCTAATGAGATGATCCGGGACCTG 396  
QY 564 CTGGAACCTCTTCCGGAGACCTGGTAATCCGAGAGACTGCGGGGGAAATATCCTGATT 623  
DB 397 CTGAACCCCTCCCTGGGCTACCTGGAGCTGCGGGAGGACTCTAAGGGGGTGATCCAGGTG 456  
QY 624 CCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTTGGCGGCACTTCCTGCCA 683  
DB 457 GCCGGCATCACCGAANN 516  
QY 684 GCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACAGCGCTCTCCCGCAGTCAAT 743  
DB 517 GGGNAACGGGAGAGGAGCCAGGAGCCACCGGCCCAACAGAGCTCTCTCCGCTCCAC 576  
QY 744 GCTGTGCTCTGTGTCAGGTGAGACGAGCGGGAACGTTTGGCCCCATTTTCGCCAGCAGAG 803  
DB 577 GCGGGCGGCTGTTC----- 592  
QY 804 GGAATACTTACCTGATTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 863  
DB 593 -----TGATCGACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 639  
QY 864 GGCCTTCGGTAAAGAGAGTGGAGCCATCAACACTCTCCCTGTTGTCTGGGCAAGTG 923  
DB 640 GGGCAGGCTATGAGGAGGGGGCCCATCAACCGTCACTGCTGCACTGGGCAACTGC 699  
QY 924 GTAGATGCGCTGAATCAG-----GGCCTCCCTGCTGTACCTTATCGGACAGCAAGCTC 977

Db 700 ATCAACGCCCTGAGCGACAAGGGTAGCAACAAGTACATCAACTATCGCGACAGCAAGCTC 759  
QY 978 ACTCGGCTATTGCAAGACTCTCTGGGTGGCTAGCCCAAGTATCCTTATTGCAACATT 1037  
Db 760 ACCCGGCTCCTGAAGGACTCTCTGGGAGGAAACAGCGCCACAGTGATCGCTCACATC 819  
QY 1038 GCCCCTGAGAGACGCTTCTACCTAGACACAGTCTCCGCACTCAACTTTGCTGCCAGTCC 1097  
Db 820 AGTCCTGCGAGCAGTGCCTTCGAGGAGTCCCGGAACACCTTGACCTACGCCGCGCGGCC 879  
QY 1098 AAGGAGGTGATCAATCGG 1115  
Db 880 AAGAACATTAAAGACTAGG 897

Search completed: November 10, 2004, 18:25:57  
Job time : 144.547 secs

	Query Match	100.0%;	Score 1538;	DB 17;	Length 1538;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1538;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	ATGCGAGCCGCGCGGCTCGACGACAGAGGCGACGCGAGATGGCGGACGCTTCACGG	60		
Db	1	ATGCGAGCCGCGCGGCTCGACGACAGAGGCGACGCGAGATGGCGGACGCTTCACGG	60		
QY	61	CGCGCATCTCAGAGAGCTGTCGCTGTCGGCTTAAGCAAGATTGGAGCTACTCGTCGTCAC	120		
Db	61	CGCGCATCTCAGAGAGCTGTCGCTGTCGGCTTAAGCAAGATTGGAGCTACTCGTCGTCAC	120		

QY 121 CTCAGCTCCGTAAGGTCGTGCGACTGCGGCCATTTGTGATGGAACAGCGGAG 180  
 DB 121 CTCAGCTCCGTAAGGTCGTGCGACTGCGGCCATTTGTGATGGAACAGCGGAG 180  
 QY 181 CAAGTGATCCCGCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240  
 DB 181 CAAGTGATCCCGCTGTGTGCGGGGATGACAGCTGCTCTCTAGAGATTGCTAACTGGA 240  
 QY 241 GGAACCAACAGAGACTCTCAAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTC 300  
 DB 241 GGAACCAACAGAGACTCTCAAAATACCAAGTTGATGCTTCTATGGGAGAGGAGTACTC 300  
 QY 301 AGCAGGACATCTAGCAGTTTCAAGTTCAGTCCAGCCATCTTAAGGCACTTGTCTGGAAGGCGA 360  
 DB 301 AGCAGGACATCTAGCAGTTTCAAGTTCAGTCCAGCCATCTTAAGGCACTTGTCTGGAAGGCGA 360  
 QY 361 ATGCCAGTGTGCTTGTGCTATGAGGCTTCAAGTTCAGTCCAGCCATCTTAAGGCACTTGTCTGGAAGGCGA 420  
 DB 361 ATGCCAGTGTGCTTGTGCTATGAGGCTTCAAGTTCAGTCCAGCCATCTTAAGGCACTTGTCTGGAAGGCGA 420  
 QY 421 GCCCAGAGCAACCTGGGGTATCCCGGGGCTCTCATGAGACCTCTGAGCTCACAAGGG 480  
 DB 421 GCCCAGAGCAACCTGGGGTATCCCGGGGCTCTCATGAGACCTCTGAGCTCACAAGGG 480  
 QY 481 AGGAGGTGCGGAGGCGCGCCATGGCCCTTCTGTCAACATGCTTACCTAGAGATCT 540  
 DB 481 AGGAGGTGCGGAGGCGCGCCATGGCCCTTCTGTCAACATGCTTACCTAGAGATCT 540  
 QY 541 ACCAGAGAGAGGATTAAGACCTCTCTGACCCCTGCTTCGGGAGACCTGGTAATCCGGAAG 600  
 DB 541 ACCAGAGAGAGGATTAAGACCTCTCTGACCCCTGCTTCGGGAGACCTGGTAATCCGGAAG 600  
 QY 601 ACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGAGCCATCAGTAGCTTTGCTG 660  
 DB 601 ACTGCGGGGGAATATCTGATTCGGGGTCTCTCCAGAGAGCCATCAGTAGCTTTGCTG 660  
 QY 661 ATTTGAGCGGCACTTCTCCAGCAGTGGAAATCGGACTGAGAGCCACCCGGCTCA 720  
 DB 661 ATTTGAGCGGCACTTCTCCAGCAGTGGAAATCGGACTGAGAGCCACCCGGCTCA 720  
 QY 721 ACCAGCGCTCTCCCGCAGTCACTGCTGTCTCTGCTCAAGTGGACCAAGCGGAAAGTT 780  
 DB 721 ACCAGCGCTCTCCCGCAGTCACTGCTGTCTCTGCTCAAGTGGACCAAGCGGAAAGTT 780  
 QY 781 TGGCCCATTTCCGACGAGAGGAAATCTTACCTGATTTGCTGGCTGGCTCAGAG 840  
 DB 781 TGGCCCATTTCCGACGAGAGGAAATCTTACCTGATTTGCTGGCTGGCTCAGAG 840  
 QY 841 ACAACCGGCGCACAGGCAACAAGGCTTCCGCTTAAAGAGAGTGGAGCCATCAACACCT 900  
 DB 841 ACAACCGGCGCACAGGCAACAAGGCTTCCGCTTAAAGAGAGTGGAGCCATCAACACCT 900  
 QY 901 CCCTGTTTGTCTGGGCAAGTGTAGATCGCTGTAATCAGGCGCTCCCTCGTGTACCTT 960  
 DB 901 CCCTGTTTGTCTGGGCAAGTGTAGATCGCTGTAATCAGGCGCTCCCTCGTGTACCTT 960  
 QY 961 ATCGGACAGAGCTCACTCGCTTATGAGAGCTCTTGGTGGCTCAGCCACAGTA 1020  
 DB 961 ATCGGACAGAGCTCACTCGCTTATGAGAGCTCTTGGTGGCTCAGCCACAGTA 1020  
 QY 1021 TCCTTATTTGCCAATTTGCCCTTGGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
 DB 1021 TCCTTATTTGCCAATTTGCCCTTGGAGAGCTTCTACCTAGACACAGTCTCCGCACTCA 1080  
 QY 1081 ACTTTGCTGCGAGTCCAAGAGTGTCAATCGGCTTTTACCAATGAGAGCTCGCAG 1140  
 DB 1081 ACTTTGCTGCGAGTCCAAGAGTGTCAATCGGCTTTTACCAATGAGAGCTCGCAG 1140  
 QY 1141 CTATGCTTTGGGACCTTTAAGTGTCTCAGAAAGAAATTTGCTTGGTCCACAGAGCAA 1200  
 DB 1141 CTATGCTTTGGGACCTTTAAGTGTCTCAGAAAGAAATTTGCTTGGTCCACAGAGCAA 1200  
 QY 1201 AGAGAGCCCGAGGCGCTGAGAGAGGAGATTTGGAGCCCTTGAGCCCTGAGCTCCAG 1260

DB 1201 AGNAGCCCGAGGCTTGGAGAGAGGATTTGGAGCCCTGAGCCATGGCAGCTCCAG 1260  
 QY 1261 CCTCTGCTCCAGAAATCTAGCCCCCTTACAGAGCTTAAGCAGATGAGCCCGCCATGC 1320  
 DB 1261 CCTCTGCTCCAGAAATCTAGCCCCCTTACAGAGCTTAAGCAGATGAGCCCGCCATGC 1320  
 QY 1321 TGGAGCGCTCTCTAGCTTGGACCCCTGCTTGCCTCCAGGGGAGCCAGGGGGCCCTC 1380  
 DB 1321 TGGAGCGCTCTCTAGCTTGGACCCCTGCTTGCCTCCAGGGGAGCCAGGGGGCCCTC 1380  
 QY 1381 TGTGTAGTACCCCAAGCGAGCGGATGCTGTCTAATGAAGACAGTAGAAGAGAAGACC 1440  
 DB 1381 TGTGTAGTACCCCAAGCGAGCGGATGCTGTCTAATGAAGACAGTAGAAGAGAAGACC 1440  
 QY 1441 TAGAGATTGAGAGCTTAAAGCAAGCAAAAGAACTGGAGGCAAGATGTTGGCCCGA 1500  
 DB 1441 TAGAGATTGAGAGCTTAAAGCAAGCAAAAGAACTGGAGGCAAGATGTTGGCCCGA 1500  
 QY 1501 AGGCTGAGAAAGAGGAAACCATTTGCCACATGTGA 1538  
 DB 1501 AGGCTGAGAAAGAGGAAACCATTTGCCACATGTGA 1538

RESULT 2  
 US-10-334-143-102  
 ; Sequence 102, Application US/10334143  
 ; Publication No. US20040009549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
 ; APPLICANT: SUDARSANAM, SUCHA  
 ; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
 ; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
 ; FILE REFERENCE: 038602/1543  
 ; CURRENT APPLICATION NUMBER: US/10/334,143  
 ; PRIOR FILING DATE: 2002-12-31  
 ; PRIOR FILING DATE: 2001-12-31  
 ; PRIOR FILING DATE: 2001-12-31  
 ; NUMBER OF SEQ ID NOS: 207  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 102  
 ; LENGTH: 2099  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-334-143-102

Query Match 98.8%; Score 1519; DB 16; Length 2099;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1530; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 6 GCCCGCGCGGCTCGACGCGAGAGGCGACGCGAGATGGCGGAGCTTCAGCGCGGCG 65  
 DB 29 GCCCGCGCGGCTCGACGCGAGAGGCGACGCGAGATGGCGGAGCTTCAGCGCGGCG 88  
 QY 66 ATCTCAGGAGCTGCTCGCTGCTGCTTAAAGCAAGATTGGAGCTACTCGTCTCCACCTCCA 125  
 DB 89 ATCTCAGGAGCTGCTCGCTGCTGCTTAAAGCAAGATTGGAGCTACTCGTCTCCACCTCCA 148  
 QY 126 GCTCGCGTAAAGGTGCTGTGCGACTGCGGCCATTTGTGATGGAACAGCGGAGCAAGT 185  
 DB 149 GCTCGCGTAAAGGTGCTGTGCGACTGCGGCCATTTGTGATGGAACAGCGGAGCAAGT 208  
 QY 186 GATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTAGAGATTGCTAATCTGAGGAAC 245  
 DB 209 GATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTAGAGATTGCTAATCTGAGGAAC 268  
 QY 246 CACCAGAGACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 305  
 DB 269 CACCAGAGACTCTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAG 328  
 QY 306 GACATCTATGAGGTTTCAAGTGTGAGCCATCTTGAAGCACTTGTGGAAGGCAAGATGCC 365  
 DB 329 GACATCTATGAGGTTTCAAGTGTGAGCCATCTTGAAGCACTTGTGGAAGGCAAGATGCC 388

```
Qy 366 AGTGTGCTTGGCTATGACCCACAGAGCTGGGAAGAGCGCACAAATGCTGGGAGGCCA 425
Db 369 AGTGTGCTTGGCTATGACCCACAGAGCTGGGAAGAGCGCACAAATGCTGGGAGGCCA 448
Qy 426 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGACCTCTGCGAGCTCAAGGGAGAG 485
Db 449 GAGCAACCTGGGGTGATCCCGGGGCTCTCATGACCTCTGCGAGCTCAAGGGAGAG 508
Qy 486 GGTGCCAGGGCCCGCCATGGCCCTTCTGTGCAATGCTTACCTAGAGATCTACAG 545
Db 509 GGTGCCAGGGCCCGCCATGGCCCTTCTGTGCAATGCTTACCTAGAGATCTACAG 568
Qy 546 GAGAAGGTATTAGACCTCTCGACCTCTCTCGGAGACCTGGTAATCCGAGAAGACTGC 605
Db 569 GAGAAGGTATTAGACCTCTCGACCTCTCTCGGAGACCTGGTAATCCGAGAAGACTGC 628
Qy 606 CGGGGGAATATCTGATTTCCGGGCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTT 665
Db 629 CGGGGGAATATCTGATTTCCGGGCTCTCTCCAGAGGCCCATCAGTAGCTTTGCTGATTT 688
Qy 666 GAGCGGCACTTCTCGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAG 725
Db 689 GAGCGGCACTTCTCGCCAGCCAGTCGAATCGGACTGTAGAGCCACCCGGCTCAACAG 748
Qy 726 CGCTCCTCCCGCAGTCATGCTGTCTCCTGCTCAAGGTGACCGGGAACTTTGGCC 785
Db 749 CGCTCCTCCCGCAGTCATGCTGTCTCCTGCTCAAGGTGACCGGGAACTTTGGCC 808
Qy 786 CCATTTGCCAGCGAGAGGAAACTCTTACCTGATTTGACTTTGGCTGGGTACAGAGCAAC 845
Db 809 CCATTTGCCAGCGAGAGGAAACTCTTACCTGATTTGACTTTGGCTGGGTACAGAGCAAC 868
Qy 846 CGGCGCAGACCAACAGGGCTTCGGCTAAAGAGAGTGAGGACATCAACACCTCCCTG 905
Db 869 CGGCGCAGACCAACAGGGCTTCGGCTAAAGAGAGTGAGGACATCAACACCTCCCTG 928
Qy 906 TTTGTCTCGGCAAGGTGATGCTGCTGAATCAGGGCCCTCCCTCGTGTACCTTTATCGG 965
Db 929 TTTGTCTCGGCAAGGTGATGCTGCTGAATCAGGGCCCTCCCTCGTGTACCTTTATCGG 988
Qy 966 GACAGCAAGCTCACTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCCTT 1025
Db 989 GACAGCAAGCTCACTCGCTATTGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCCTT 1048
Qy 1026 ATTGCCAATGTCCTCGAGAGCGCTTCTACCTAGACAGCTCGGCACTCAACTTT 1085
Db 1049 ATTGCCAATGTCCTCGAGAGCGCTTCTACCTAGACAGCTCGGCACTCAACTTT 1108
Qy 1086 GCTGCCAGGTCACAGGAGGTGATCAATCGGCCTTTTACCAATGAGAGCTCGACCTCAT 1145
Db 1109 GCTGCCAGGTCACAGGAGGTGATCAATCGGCCTTTTACCAATGAGAGCTCGACCTCAT 1168
Qy 1146 GCCTTGGGACCTGTTAAGCTCTCAGAAAGATTGCTTGTGTCACAGAGGCAAGAGA 1205
Db 1169 GCCTTGGGACCTGTTAAGCTCTCAGAAAGATTGCTTGTGTCACAGAGGCAAGAGA 1228
Qy 1206 GCGCGAGGCTGAGGAGAGGATTTGGAGCCCTGAGCCCATGCGAGCTCCAGCCTCT 1265
Db 1229 GCGCGAGGCTGAGGAGAGGATTTGGAGCCCTGAGCCCATGCGAGCTCCAGCCTCT 1288
Qy 1266 GCCTTCCAGAACTCAGCCCCCTACAGAAAGCTAAGAGCAGCATGGACCCGGCCATGCTGGAG 1325
Db 1289 GCCTTCCAGAACTCAGCCCCCTACAGAAAGCTAAGAGCAGCATGGACCCGGCCATGCTGGAG 1348
Qy 1326 CGCTCTCTCAGCTTGGACCGTCTGCTGCTCCAGGGGAGCCAGGGGGCCCCCTCTGTTG 1385
Db 1349 CGCTCTCTCAGCTTGGACCGTCTGCTGCTCCAGGGGAGCCAGGGGGCCCCCTCTGTTG 1408
Qy 1386 AGTACCCCAAGCCGAGCGGATGGTGTATGAGACAGTAGAGAGGAGGACTAG-A 1444
Db 1409 AGTACCCCAAGCCGAGCGGATGGTGTATGAGACAGTAGAGAGGAGGACTAGTA 1468
```

```
Qy 1445 GATTGAGAGCTTAAGACGAAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGGC 1504
Db 1469 GATTGAGAGCTTAAGACGAAAGCAAAAGAACTGGAGGCCAAGATGTTGGCCCAAGGC 1528
Qy 1505 TGAGAAAGAGGAACCATTTGTCACCAATG 1535
Db 1529 TGAGAAAGAGGAACCATTTGTCACCAATG 1559

RESULT 3
US-10-797-893-7
; Sequence 7, Application US/10797893
; Publication No. US20040142397A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1044
; CURRENT FILING DATE: 2004-03-09
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/724,224
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/597,292
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Human
US-10-797-893-7

Query Match 95.0%; Score 1461; DB 17; Length 1464;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAAG 137
Db 4 GGTGCTGTGCGCTTAAGCAAGATTGGAGCTACTCGTGTCCACCTCCAGCTCGGTAAAG 63
Qy 138 GTGCTGTGCGACTGCGGCCCATTTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 197
Db 64 GTGCTGTGCGACTGCGGCCCATTTGTGGATGGAAACAGCGGAGCAAGTGATCCCCCTGT 123
Qy 198 GTGCGGGCATGGACAGCTGCTCTAGAGATTCTTAACCTGGAGGAAACCAACAGAGACT 257
Db 124 GTGCGGGCATGGACAGCTGCTCTAGAGATTCTTAACCTGGAGGAAACCAACAGAGACT 183
Qy 258 CTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317
Db 184 CTCAAAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243
Qy 318 GGTTCAGTGACGCCCATCTTAAGCACTTGTGTGAAGGCGCAGAAATGCCAGTGTGCTGCC 377
Db 244 GGTTCAGTGACGCCCATCTTAAGCACTTGTGTGAAGGCGCAGAAATGCCAGTGTGCTGCC 303
Qy 378 TATGAGCCCAACAGAGCTGGGAAGACCCACACAAATGCTGGCAGCCCAACCACTGGG 437
Db 304 TATGAGCCCAACAGAGCTGGGAAGACCCACACAAATGCTGGCAGCCCAACCACTGGG 363
Qy 438 GTGATCCCGGGGCTCTCATGGACCTCTCATGGAGCTCAAGGGAGAGGGTGGCCAGGGGC 497
Db 364 GTGATCCCGGGGCTCTCATGGAGCTCTCATGGAGCTCAAGGGAGAGGGTGGCCAGGGGC 423
Qy 498 CGGCGATGGGCCCTTTCTGTCAACATGCTTTA CCTAGAGATCTTACAGGAGAAAGTATTA 557
Db 424 CGGCGATGGGCCCTTTCTGTCAACATGCTTTA CCTAGAGATCTTACAGGAGAAAGTATTA 483
Qy 558 GACCTCTCGACCTGCTTGGGAGACTGCTGTAATCCGAGAAGACTCCCGGGGGAATATC 617
Db 484 GACCTCTCGACCTGCTTGGGAGACTGCTGTAATCCGAGAAGACTCCCGGGGGAATATC 543
Qy 618 CTGATTCGGGCTCTCTCCCAAGAGCCCATCAGTAGCTTTGTGATTTTGGCGGCACCTTC 677
```







QY 961 ATCGGACAGCAAGCTCACTCGCCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTA 1020  
 Db 961 ATCGGACAGCAAGCTCACTCGCCTATTGCGAGACTCTCTGGGTGGCTCAGCCACAGTA 1020  
 QY 1021 TCCTTATTGCCAATTTGCCCTTGAGAGAGCGTTCTTACCTAGACACAGTCTCCGCACTCA 1080  
 Db 1021 TCCTTATTGCCAATTTGCCCTTGAGAGAGCGTTCTTACCTAGACACAGTCTCCGCACTCA 1080  
 QY 1081 ACTTTGCTGCAGGTCCAAGGAGTGATCAATCG 1114  
 Db 1081 ACTTTGCTGCAGGTCCAAGGAGTGATCAATCG 1114

RESULT 6

US-10-797-893-5  
 ; Sequence 5, Application US/10797893  
 ; Publication No. US20040142397A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; their use  
 ; FILE REFERENCE: 1044  
 ; CURRENT APPLICATION NUMBER: US/10/797,893  
 ; CURRENT FILING DATE: 2004-03-09  
 ; PRIOR APPLICATION NUMBER: US/09/724,224  
 ; PRIOR FILING DATE: 2000-11-28  
 ; PRIOR APPLICATION NUMBER: 09/597,292  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 1041  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-10-797-893-5

Query Match 67.3%; Score 1035.4; DB 17; Length 1041;  
 Best Local Similarity 99.9%; Pred. No. 1.5e-306;  
 Matches 1036; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 78 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGCGTAAGG 137  
 Db 4 GGTGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGCGTAAGG 63  
 QY 138 GTGGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGCGTAAGG 197  
 Db 64 GTGGCTGTGCGCTAAGCAAGATTGGAGTACTCGTCTCCACCTCCAGCTCGCGTAAGG 123  
 QY 198 GTGGGGGCAATGGACAGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 257  
 Db 124 GTGGGGGCAATGGACAGCTCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACT 193  
 QY 258 CTCAATACCAAGTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 317  
 Db 184 CTCAATACCAAGTTGATGCTCTTATGGGAGAGGAGTACTCAGCAGGACATCTATGCA 243  
 QY 318 GGTTCAGTGCAGCCCATCTTAAGGCATCTCTGAGAGGGCAGAAATGCCAGTGTCTTGGC 377  
 Db 244 GGTTCAGTGCAGCCCATCTTAAGGCATCTCTGAGAGGGCAGAAATGCCAGTGTCTTGGC 303  
 QY 378 TATGGACCCACAGAGCTGGGAAGACGACACAAATGCTGGGAGCCAGAGCAACTGGG 437  
 Db 304 TATGGACCCACAGAGCTGGGAAGACGACACAAATGCTGGGAGCCAGAGCAACTGGG 363  
 QY 438 GTGATCCCGGGCTCTCATGAGACCTCTCAGCTCACAAGGAGGAGGCTGCCGAGGGC 497  
 Db 364 GTGATCCCGGGCTCTCATGAGACCTCTCAGCTCACAAGGAGGAGGCTGCCGAGGGC 423  
 QY 498 CGGCCATGGGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTTACAGGAGGATTA 557  
 Db 424 CGGCCATGGGCCCTTTCTGTACCATGCTCTTACCTAGAGATCTTACAGGAGGATTA 483

QY 558 GACCTCTCGAACCCTCTTCGGGAGACCTGGTAAATCCGAGAGACTGCGCGGGGAATATC 617  
 Db 484 GACCTCTCGAACCCTCTTCGGGAGACCTGGTAAATCCGAGAGACTGCGCGGGGAATATC 543  
 QY 618 CTGATTCGGGGTCTCTCCCAAGAGCCCATCAGTAGCTTTTGTGATTTTGTAGCGGCACTTC 677  
 Db 544 CTGATTCGGGGTCTCTCCCAAGAGCCCATCAGTAGCTTTTGTGATTTTGTAGCGGCACTTC 603  
 QY 678 CTGCGACAGTGGAAATCGGACTGTAGAGCCACCCGGCTCAACCGCCCTCTCTCCGC 737  
 Db 604 CTGCGACAGTGGAAATCGGACTGTAGAGCCACCCGGCTCAACCGCCCTCTCTCCGC 663  
 QY 738 AGTCATGCTGTCTCTCTGGTCAAGGTGGACCGGGAACGTTTGGCCCCATTTTCGCCAG 797  
 Db 664 AGTCATGCTGTCTCTCTGGTCAAGGTGGACCGGGAACGTTTGGCCCCATTTTCGCCAG 723  
 QY 798 CGAGAGGGAAAACTCTACCTGATTGATTGGCTGGGTCAAGAGCAACCGCGGCACAGGC 857  
 Db 724 CGAGAGGGAAAACTCTACCTGATTGATTGGCTGGGTCAAGAGCAACCGCGGCACAGGC 783  
 QY 858 ACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTTGTTCCTGGGC 917  
 Db 784 ACAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTTGTTCCTGGGC 843  
 QY 918 AAGTGTGTAGATCGCTGATCAATCAGGGCTCCCTCTGTTGTTTATCGGACAGCAAGCTC 977  
 Db 844 AAGTGTGTAGATCGCTGATCAATCAGGGCTCCCTCTGTTGTTTATCGGACAGCAAGCTC 903  
 QY 978 ACTCGCTATTGACAGACTCTCTGGTGGTCAAGCCACAGTATCTTATTTGCAACATT 1037  
 Db 904 ACTCGCTATTGACAGACTCTCTGGTGGTCAAGCCACAGTATCTTATTTGCAACATT 963  
 QY 1038 GCGCCCTGAGAGAGCTTCTACCTAGACAGTCTCGCAGTCAACTTGTGTCAGGTC 1097  
 Db 964 GCGCCCTGAGAGAGCTTCTACCTAGACAGTCTCGCAGTCAACTTGTGTCAGGTC 1023  
 QY 1098 AAGGAGGTGATCAATCG 1114  
 Db 1024 AAGGAGGTGATCAATCG 1040

RESULT 7

US-09-960-253-131  
 ; Sequence 131, Application US/09960253  
 ; Patent No. US20020123619A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Benson, Darin R.  
 ; APPLICANT: Mohamath, Raodoh  
 ; APPLICANT: Lodes, Michael J.  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.556  
 ; CURRENT APPLICATION NUMBER: US/09/960,253  
 ; CURRENT FILING DATE: 2001-09-20  
 ; NUMBER OF SEQ ID NOS: 187  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 131  
 ; LENGTH: 491  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-960-253-131

Query Match 31.2%; Score 480; DB 9; Length 491;  
 Best Local Similarity 99.8%; Pred. No. 2.1e-136;  
 Matches 491; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1010 AGCCACAGTATCTTATTTGCCAAATTTGCCCTTGAGAGACGCTTCTACTAGACAGT 1069  
 Db 1 AGCCACAGTATCTTATTTGCCAAATTTGCCCTTGAGAGACGCTTCTACTAGACAGT 60  
 QY 1070 CTCGCGACTCAACTTTGCTGCCAGGTCCCAAGGAGGTGATCAATCGGCCCTTTTACCAATGA 1129  
 Db 61 CTCGCGACTCAACTTTGCTGCCAGGTCCCAAGGAGGTGATCAATCGGCCCTTTT-CCAATGA 119



```
QY 1130 GAGCCTGCAGCCTCATGCTTTGGACCTGTTAAGCTGTCTCAGAAAGAAATTGCTTGTCC 1189
Db 120 GAGCCTGCAGCCTCATGCTTTGGACCTGTTAAGCTGTCTCAGAAAGAAATTGCTTGTCC 179
QY 1190 ACCGAGCAAGAGAGAGCCGAGGCTTGAAGAGAGAGAGATTGGAGCCCTTGAGCCCAT 1249
Db 180 ACCGAGCAAGAGAGAGCCGAGGCTTGAAGAGAGAGAGATTGGAGCCCTTGAGCCCAT 239
QY 1250 GGCAGCTCCAGCTCTGCTCTCCAGAACTCAGCCCTCAGAGCTTAAGCAGCATGGA 1309
Db 240 GGCAGCTCCAGCTCTGCTCTCCAGAACTCAGCCCTCAGAGCTTAAGCAGCATGGA 299
QY 1310 CCGGCCATGTGGAGCGCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGGAGCCA 1369
Db 300 CCGGCCATGTGGAGCGCTCTCAGCTTGGACCGTCTGCTTGCCTCCAGGGGAGCCA 359
QY 1370 GGGGGCCCTCTGTTAGTACCCCAAGCGAGAGCGGATGCTTAATGAAGACAGTAGA 1429
Db 360 GGGGGCCCTCTGTTAGTACCCCAAGCGAGAGCGGATGCTTAATGAAGACAGTAGA 419
QY 1430 AGAAGAGACCTAGAGATTGAGAGCTTAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 1489
Db 420 AGAAGAGACCTAGAGATTGAGAGCTTAAGACGAAGCAAAAGAACTGGAGGCCAAGAT 479
QY 1490 GTTGGCCCGAGAA 1501
Db 480 GTTGGCCCGAGAA 491
```

## RESULT 8

US-09-803-719-1937

; Sequence 1937, Application US/09803719

; Publication No. US2003004783A1

; GENERAL INFORMATION:

; APPLICANT: Williams, Lewis T.

; APPLICANT: Escobedo, Jaime

; APPLICANT: Innis, Michael A.

; APPLICANT: Garcia, Pablo Dominiguez

; APPLICANT: Sudduth-Klinger, Julie

; APPLICANT: Reinhard, Christoph

; APPLICANT: Giese, Klaus

; APPLICANT: Randazzo, Filippo

; APPLICANT: Kennedy, Giulia C.

; APPLICANT: Pot, David

; APPLICANT: Kassam, Altaf

; APPLICANT: Lamson, George

; APPLICANT: Drmanac, Radoje

; APPLICANT: Crkvenjakov, Radomir

; APPLICANT: Dickson, Mark

; APPLICANT: Drmanac, Snezana

; APPLICANT: Labat, Ivan

; APPLICANT: Leshkowitz, Dena

; APPLICANT: Kita, David

; APPLICANT: Garcia, Veronica

; APPLICANT: Jones, Lee William

; APPLICANT: Stache-Crain, Birgit

; TITLE OF INVENTION: Human Genes and Gene Products

; FILE REFERENCE: 1624.002

; CURRENT APPLICATION NUMBER: US/09/803,719

; CURRENT FILING DATE: 2001-03-09

; PRIOR APPLICATION NUMBER: 60/188,609

; PRIOR FILING DATE: 2000-03-09

; NUMBER OF SEQ ID NOS: 2396

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1937

; LENGTH: 386

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-803-719-1937

Query Match 24.5%; Score 376.4; DB 10; Length 386;

Best Local Similarity 98.4%; Pred. No. 1.1e-104;

```
Matches 380; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1147 CTTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATTGCTTGTCCACCCAGAGCAAGAGAG 1206
Db 1 CGTTGCTGTGCTTAAGCTGTCTCAGAAAGAAATTGCTTGGTCCACCCAGAGCAAGAGAG 60
QY 1207 CCGGAGCCCTTGAGAAAGAGAGATTGGAGCCCTTGAGCCCATGCGACCTCCAGCCTCTG 1266
Db 61 CCGGAGCCCTTGAGAAAGAGAGATTGGAGCCCTTGAGCCCATGCGACCTCCAGCCTCTG 120
QY 1267 CTTCCGAAACTCAGCCCTTACAGAACTTAAGCAGCATGAGCCGCGCATGCTGGAGC 1326
Db 121 CTTCCGAAACTCAGCCCTTACAGAACTTAAGCAGCATGAGCCGCGCATGCTGGAGC 180
QY 1327 GCCTCCTCAGCTTGGACCGTCTGCTCTCCAGGGGAGCCAGGGGCCCTCTGTTGA 1386
Db 181 GCCTCCTCAGCTTGGACCGTCTGCTCTCCAGGGGAGCCAGGGGCCCTCTGTTGA 240
QY 1387 GTACCCCAAGCGAGAGCGGATGCTTAATGAAGACAGTAGAAGAGACCTTAGAGA 1446
Db 241 GTACCCCAAGCGAGAGCGGATGCTTAATGAAGACAGTAGAAGAGACCTTAGAGA 300
QY 1447 TTGAGAGCTTAAGCAGCAAGCAAAAGAACTGGAGGGCCCAAGATGTTGCCCAAGAGCTG 1506
Db 301 TTGAGAGCTTAAGCAGCAAGCAAAAGAACTGGAGGGCCCAAGATGTTGCCCAAGAGCTG 360
QY 1507 AGGAAGAGGAGAACCATTTGCCACA 1532
Db 361 AGGAAGAGGAGAACCATTTGCCACA 386
```

## RESULT 9

US-10-087-192-1180

; Sequence 1180, Application US/10087192

; Publication No. US20020182586A1

; GENERAL INFORMATION:

; APPLICANT: Morris, David W.

; APPLICANT: Engelhard, Eric K.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1180

; LENGTH: 24525

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-087-192-1180

Query Match 23.7%; Score 365; DB 13; Length 24525;

Best Local Similarity 73.8%; Pred. No. 1.4e-100;

Matches 597; Conservative 0; Mismatches 0; Indels 212; Gaps 2;

```
QY 550 AGGTATTAGACCTCCTCGAGCCCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGG 609
Db 2313 AGGTATTAGACCTCCTCGAGCCCTGCTTCGGGAGACCTGGTAATCCGAGAGACTGCCGGG 2372
QY 610 GGAATATCCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGC 669
Db 2373 GGAATATCCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTGAGC 2432
QY 670 GGCACCTTCTGCCAGCCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCT 729
Db 2433 GGCACCTTCTGCCAGCCAGTCGAAATCGGACTGTAGAGCCACCCGGCTCAACAGCGCT 2492
QY 730 CCTCCCGCAGTCAGCTGCTGCTCCTCGTCA----- 759
```

Db 2493 CCTCCGACATCATGTGTCTCTCTGCTCAAGGTGAGGCGCGAGACAGAGGCGCGAGGACCT 2552  
 QY 760 -----AGGTGGACCG 770  
 Db 2553 GGAAGCCAGGAGGCTGAGCTAAGCAGAGACCTTTGTTCTTACCCCCAGGTGGACCG 2612  
 QY 771 CGGGAAGCTTTGGCCCCCATTTCCGACGAGAGGGAACCTTACCTGATTGACTTTGGCT 830  
 Db 2613 CGGGAAGCTTTGGCCCCCATTTCCGACGAGAGGGAACCTTACCTGATTGACTTTGGCT 2672  
 QY 831 GGCTCAGAGACAAACCGCGCCACAGGCAACAAGGGCTTCGGCTAAAGAGAGTGAAGCC 890  
 Db 2673 GGCTCAGAGACAAACCGCGCCACAGGCAACAAGGGCTTCGGCTAAAGAGAGTGAAGCC 2732  
 QY 891 ATCAACACCTCCCTGTTGTCCTGGCCAAAGTGGTAGATCGCTGAATCAGGGCCCTCCCT 950  
 Db 2733 ATCAACACCTCCCTGTTGTCCTGGCCAAAGTGGTAGATCGCTGAATCAGGGCCCTCCCT 2792  
 QY 951 CGTGTACCTTATCGGACAGCAAGCTCACTCGCTATTG----- 989  
 Db 2793 CGTGTACCTTATCGGACAGCAAGCTCACTCGCTATTGCGCTATTGCGCTCAGGCCACCTGTCTC 2852  
 QY 990 ----- 999  
 Db 2853 AGGGAAGAGGGGCTGCAGAAAGAGGTTCTCAGGCCCTGCTGTGGGTGGGGAATAGAGT 2912  
 QY 990 -----CAGGACTC 997  
 Db 2913 TGAGGCATAGGAAGGTGGGCTTCTGACCCACCACTGCGTGTCTCACTCAGGACTC 2972  
 QY 998 TCTGGGTGGCTCAGCCACAGTATCTTATTGCAACATTTGCCCTTGAGAGAGCTTCTTA 1057  
 Db 2973 TCTGGGTGGCTCAGCCACAGTATCTTATTGCAACATTTGCCCTTGAGAGAGCTTCTTA 3032  
 QY 1058 CTTAGACACAGTCTCCGCACTCACTTATGCTGCGAGTCAAGAGGTGATCAATCGGCC 1117  
 Db 3033 CTTAGACACAGTCTCCGCACTCACTTATGCTGCGAGTCAAGAGGTGATCAATCGGCC 3092  
 QY 1118 TTTTACCATGAGAGCTCGAGCTCATG 1146  
 Db 3093 TTTTACCATGAGAGCTCGAGCTCATG 3121

RESULT 10  
 US-09-918-995-31705  
 ; Sequence 31705, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 ; FROM VARIOUS CDNA LIBRARIES  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 31705  
 ; LENGTH: 464  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(464)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-918-995-31705

Query Match 17.2%; Score 265; DB 10; Length 464;  
 Best Local Similarity 77.8%; Pred. No. 1.7e-70;  
 Matches 333; Conservative 0; Mismatches 95; Indels 1; Gaps 1;  
 QY 200 GCGGGCATGGACAGCTGTCTCTAGAGATTGCTAACTGGAGGAACCAACGAGGAGACTCT 259

Db 37 GCCTGGAATTCATATATCGCGTGCATAGAGTGTAAACGCGCCGAAACCAACCCGATAATCT 96  
 QY 260 CAATACCAAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGAGG 319  
 Db 97 CCCTTACCTGTGCTTTGCTTTTACAGCAGCATAGGAGGACTGCGCTTGACATGTACCTTCC 156  
 QY 320 TTCAGTGCAGCCCATCTTAAGGCACCTTGTGGAAGGCGAGATGCCAGTGTGCTTGCCTA 379  
 Db 157 TTGACCGTGGACCGTGCCTCTCTTTCTGTAATAAAGCGAGGAGCTGGCCCTA 216  
 QY 380 TGGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGCGAGCCAGAGCAACCTGGGGT 439  
 Db 217 CGGTGCTGAGGAGCTGGCATTAC-ATCTCGATGCTTGGCAGCCAGAGCAACCTGGGGT 275  
 QY 440 GATCCCGCGGCTCTCATGGACCTCTCAGCTCACAAGGGAGAGGGTGCAGGCGCG 499  
 Db 276 GATCCCGCGGCTCTCATGGACCTCTCAGCTCACAAGGGAGAGGGTGCAGGCGCG 335  
 QY 500 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCAGAGAGGATATCTCT 559  
 Db 336 GCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCAGAGAGGATATCTCT 395  
 QY 560 CTTCTTGGACCTTCTGTCGAGGAGACCTGCTTAATCCGAGAGACTGCGCGGGGAATATCTCT 619  
 Db 396 CTTCTTGGACCTTCTGTCGAGGAGACCTTGGAAATCCGAGAGACTGCGCGGGGAATATCTCT 455  
 QY 620 GATTCGGG 628  
 Db 456 GATTCGGG 464

RESULT 11  
 US-10-087-192-1177  
 ; Sequence 1177, Application US/10087192  
 ; Publication No. US20020182586A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morris, David W.  
 ; APPLICANT: Engelhard, Eric K.  
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
 ; FILE REFERENCE: 529452000122  
 ; CURRENT APPLICATION NUMBER: US/10/087,192  
 ; CURRENT FILING DATE: 2002-03-01  
 ; PRIOR APPLICATION NUMBER: US 09/747,377  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: US 09/798,586  
 ; PRIOR FILING DATE: 2001-03-02  
 ; NUMBER OF SEQ ID NOS: 2059  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1177  
 ; LENGTH: 24291  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(24291)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-087-192-1177

Query Match 17.1%; Score 262.4; DB 13; Length 24291;  
 Best Local Similarity 73.8%; Pred. No. 4e-69;  
 Matches 385; Conservative 0; Mismatches 61; Indels 76; Gaps 1;  
 QY 550 AGGTATTAGACCTTCTGGAACCTTCTCGGAGAGCTTGGTAATCCGAGAGACTGCGGG 609  
 Db 2939 AGGTATTAGACCTTCTGGAACCTTCTCGATCAGGAGACCTCTGATCCGGAAGACTGCGGAG 2998  
 QY 610 GGAATATCTGATTCGGGTCTCTCCAGAACCCATCAGTCTTGCTGATTGAGC 669  
 Db 2999 GGAACATCTGATCCAGGCTCTCACACAGAACCCATCAGTCTTGCTGACTTGAGC 3058  
 QY 670 GGCACCTTCTGCGCAGCCAGTTCGAAATCGGACTGTAGGAGCCACCGGCTCAACACGCGCT 729

Db 3059 AGCAGCTTCCTCCAGCCAGTAGAATCGAGCTGTAGGAGCCACCGGCTTAACCAAGCGT 3118  
Qy 730 CCTCCGCGAGTCATGCTGTGCTCTCTGGTCA-----759  
Db 3119 CTTCCGCTAGTACGAGTGTCTTGGTCAAGGTAAGGCCCTGACAGAAAGGCTGGG 3178  
Qy 760 -----AGTGGACAGCGG 773  
Db 3179 AAGCCCTGAATGTGGAGTGAAGTCTGCTCTCTTCTCTTCCCTAGTAGATCAGCGT 3238  
Qy 774 GAACGTTTGGCCCAATTTCGCGAGGAGGAGGAAAACCTCTACCTGATTCAGTTCGCTGGG 833  
Db 3239 GAACGTTTGAATTCATTCGCGAGGAGGAGGAAAACCTCTACCTGATTCAGTTCGCTGGT 3298  
Qy 834 TCAGAGGACAAACCGCGGACAGGCAACAAAGGSCCTTCGCTAAAGAGAGTGGAGCCATC 893  
Db 3299 TCAGAGGACAAACCGTCCGACAGGCAACAAAGGSCCTTCGCTAAAGAGAGGAGGAGCCATC 3358  
Qy 894 AACACCTCCCTGTTTGTCTGGGCAAGTGTGTAGATGGCTGAATCAGGCGCTCCCTCGT 953  
Db 3359 AACACCTCCCTGTTTGTCTGGGCAAGTGTGTAGATGGCTGAATCAGGCGCTCCCTCGC 3418  
Qy 954 GTACCTTATCGGACAGCAAGCTCACTCGCTTATTCAGGAC 995  
Db 3419 ATACCATACCGGACAGCAAGCTCACTCGTCTGCTGCAGGTC 3460

RESULT 12  
US-10-029-386-10603  
; Sequence 10603, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 10603  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7  
; OTHER INFORMATION: NT HIT: AB017334.1, EVALUE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUE 2.00e-36  
; OTHER INFORMATION: EST\_HUMAN HIT: BF975048.1, EVALUE 1.00e-126  
US-10-029-386-10603

Query Match 15.2%; Score 233.2; DB 15; Length 531;  
Best Local Similarity 98.7%; Pred. No. 9.9e-61;  
Matches 235; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 758 CAAGGTGACAGCGGGAACGTTTGGCCCCATTTCCGCGAGGAGGAGGAACTCTACCT 817  
Db 34 CAAGGTGACAGCGGGAACGTTTGGCCCCATTTCCGCGAGGAGGAGGAACTCTACCT 93  
Qy 818 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGCGTAAA 877  
Db 94 GATTGACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGCGTAAA 153  
Qy 878 AGAGAGTGGAGCCATCAACACCTCTCTTTGTCTTGGGCAAGTGTGTAGATGGCTGAA 937  
Db 154 AGAGAGTGGAGCCATCAACACCTCTCTTTGTCTTGGGCAAGTGTGTAGATGGCTGAA 213

Qy 938 TCAGGGCTCCCTCGTGTACCTTATTCGGGACAGCAAGCTCACTCGCCTATTTCAGGAC 995  
Db 214 TCAGGGCTCCCTCGTGTACCTTATTCGGGACAGCAAGCTCACTCGCCTATTTCAGGAC 271  
RESULT 13  
US-10-029-386-24303  
; Sequence 24303, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 24303  
; LENGTH: 232  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7  
; OTHER INFORMATION: SWISSPROT HIT: Q14807, EVALUE 5.00e-37  
; OTHER INFORMATION: EST\_HUMAN HIT: BE382882.1, EVALUE 1.00e-125  
; OTHER INFORMATION: NT HIT: gi6453817, EVALUE 1.00e-125  
US-10-029-386-24303

Query Match 15.0%; Score 230.4; DB 15; Length 232;  
Best Local Similarity 99.6%; Pred. No. 5.4e-60;  
Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 762 GTGACACGAGCGGGAACGTTTGGCCCCATTTCCGCGAGGAGGAGGAACTCTACCTGATT 821  
Db 1 GTGACACGAGCGGGAACGTTTGGCCCCATTTCCGCGAGGAGGAGGAACTCTACCTGATT 60  
Qy 822 GACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGGCTAAAAGAG 881  
Db 61 GACTTGGCTGGGTTCAGAGGACAAACCGGCGCACAGGCAACAGGCGCTTCGGCTAAAAGAG 120  
Qy 882 AGTGGAGCCATCAACACCTCTCTTTGTCTTGGGCAAGTGTGTAGATGGCTGAATCAG 941  
Db 121 AGTGGAGCCATCAACACCTCTCTTTGTCTTGGGCAAGTGTGTAGATGGCTGAATCAG 180  
Qy 942 GGCCTCCCTCGTGTACCTTATTCGGGACAGCAAGCTCACTCGCCTATTTCGAGG 993  
Db 181 GGCCTCCCTCGTGTACCTTATTCGGGACAGCAAGCTCACTCGCCTATTTCGAGG 232

RESULT 14  
US-10-108-260A-249  
; Sequence 249, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 249  
; LENGTH: 2095  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-249

```

Query Match      12.1%; Score 185.6; DB 16; Length 2095;
Best Local Similarity 53.7%; Pred. No. 6.1e-46;
Matches 464; Conservative 0; Mismatches 379; Indels 21; Gaps 3;

QY 264 TACCAGTTTGTATGCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCGAGTTCA 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 TACCTGTTTCAGCTGGCTTTGACTTACCGCCACCCAGGAGATGTTATCAGGCCACC 345

QY 324 GTGCAGCCCTCTTAAGGCACTTCTCTGGAAGGAGAAATGCAAGTGTGCTTGTCTATGGA 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 ACCAAGAGCCTCATCGAGGGCGTCTATCAGGCTTACAATGCCACTGTCTTTGCTATGGC 405

QY 384 CCACAGAGCTGGGAGAGCGCACACAATCTGGGAGCCAGCCAGCAACCTGGGGTATC 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 CCACAGGCTGTGGGAAACCTTACCATTCTGGGACAGACAGGAGGCTTGGCATCTAT 465

QY 444 CCGGGGCTCTCATGAGACCTCTCTGAGCTCACAGGAGGAGGCTGCCGAGGGCCGCCA 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 GTTCAGACCTTCAACGACCTTCTCCGTGCCATCGAGGAGACCATGA-----C 516

QY 504 TGGGCCCTTTCTGTCACATGCTTTACTAGATCTTACAGGAGAAAGGTATTAGACCTC 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 ATGGAGTATGAGGTCTCCATGTCTTACTGAGATCTTACAATGAGATGATCCGGGACCTG 576

QY 564 CTGACCCCTGTTCCGGGAGACCTGTAATCCGAGAGACTGCCGGGGGAATATCCTGATT 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 CTGNAACCCCTCCCTGGGCTTACTGGAGCTGGGAGGACTCTAAGGGGGGTATCCAGGTG 636

QY 624 CCGGCTCTCTCCAGAAAGCCCATCAGTAGCTTTGCTGATTTTGAAGCGGACCTTCTGCGCA 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 GCGGATCATCCGAAGTCTCCACCATCAATGCCAAGAGATCATGACAGTCTGATGAAG 696

QY 684 GCGAGTGAATCGGACTGTAGAGCCACCGGCTCAACAGCGCTCTCCCGGAGTAT 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 GGGAAACCGGAGAGCCCGAGGAGCCCGCCAGCAGCTCTCCCGTCCCGC 756

QY 744 GCTGTGCTCTGTGTCAGGTGGACCGGAGACGTTTGGCCCATTTGCGCA-----G 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 757 GCGGTACTGAGGTGACCGTGGCCGCGCAGCGCGGTCAAGAACATCTTTCAGAGGTG 816

QY 798 CGAGAGGAAATCTACTGATTGACTTGGCTGGGTGAGAGACAAACCGGCGCACAGGC 857
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 817 CGCAGGCGCCGCTGTTTCATGATCGACTGGCTGAGAGCGGCTCGCAGACACAG 876

QY 858 AACAGGCGCTTGGCTTAAAGAGAGTGGAGCCATCAACCTCCCTGTTTGTCTGGGC 917
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 877 AATGCTGGGAGCGTATGAGAGGAGGGGGCCACATCAACCGCTCACTGTGGCACTGGGC 936

QY 918 AAAGTGTGTAGATGGCTG-----AATCAGGGCTCCCTCGTGTACCTTATCGGACAGC 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 937 AACTGCATCAAGCCCTGAGGACAGGGTATGACAAAGTACATCAACTATCGGACAGC 996

QY 972 AAGTCTACTCGCCTATTGACGAGCTCTCTGGGTGGCTCAGCCCAAGTATCTTATTGCC 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 997 AAGTCTACCCGGCTCTCTGAGGACTCTCTGGAGGAAACAGCGCACAGTATGATCGCT 1056

QY 1032 AACATTCCTCTGAGAGAGCTTCTACTAGACAGCTCTCCGCACTCACTTTGTCTGCC 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1057 CACATCATGCTCTGGAGAGCTGCTTCGAGAGGTCCCGGAAACCCCTGACCTACGCGGC 1116

QY 1092 AGGTCCAAAGAGGTGATCAATCGG 1115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1117 CGGCGCAAGACATTAGACTAGG 1140

```

RESULT 15

```

US-10-287-226-301
; Sequence 301, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,

```

```

; APPLICANT: Boldog, Ference.
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: Dipippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Baha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernhet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 301
; LENGTH: 3374
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(3351)
; US-10-287-226-301

```

```

Query Match      10.2%; Score 156.8; DB 16; Length 3374;
Best Local Similarity 54.1%; Pred. No. 4.8e-37;
Matches 395; Conservative 0; Mismatches 317; Indels 18; Gaps 3;

```



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:25:02 ; Search time 5346.62 Seconds  
(without alignments)  
10482.189 Million cell updates/sec

Title: US-10-797-893-3

Perfect score: 1538

Sequence: 1 atgcagcggcggtcg.....accattgccacaatgtga 1538

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_hic.\*

4: gb\_est3.\*

5: gb\_est4.\*

6: gb\_est5.\*

7: gb\_est6.\*

8: gb\_gsl1.\*

9: gb\_gsl2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1526.8	99.3	2057	3	CR615781
2	1512.8	98.4	2025	3	CR590974
3	1462.8	95.1	2016	3	CR590954
4	1074.8	69.9	2055	3	AK084310
5	1070.6	69.6	2056	3	AK075922
6	1054.4	68.6	1124	1	AL545712
7	1002.8	65.2	1027	5	BX337406
8	980.8	63.8	1083	1	AL559772
9	978.6	63.6	1032	5	BX337353
10	942	61.2	959	5	BX363358
11	934	60.7	958	5	BX354166
12	922.4	60.0	1031	4	BM910587
13	921.4	59.9	999	1	AL559769
14	921.4	59.9	1094	5	BX376199
15	909	59.1	1019	4	BM546877
16	907.8	59.0	978	5	BX400806
17	905	58.8	1016	1	AL518361
18	899.8	57.9	1051	4	BM806996
19	889.6	57.8	980	5	BX404364
20	889.2	57.8	1041	4	BM478215
21	887.4	57.7	915	5	BX383181
22	882.8	57.4	1034	5	BQ051730
23	875.4	56.9	971	5	BQ053671
24	862.4	56.1	926	1	AL578338

25	861.8	56.0	1040.	4	BM546897
26	861.6	56.0	991	5	BX333336
27	851.8	55.4	865	1	AL555550
28	851.4	55.4	865	6	CD109092
29	844	54.9	990	5	BX399192
30	836.4	54.4	1139	4	BM553107
31	833	54.2	838	5	BX341048
32	831.8	54.1	851	5	BUI88590
33	831.6	54.1	942	5	BU527647
34	823	53.5	995	1	AL559752
35	821.8	53.4	846	5	BX404325
36	821.4	53.4	1640	3	AK088033
37	820.8	53.4	978	4	BI754589
38	819.4	53.3	981	4	BM561588
39	819	53.3	921	5	BU860011
40	816.4	53.1	878	1	AL576578
41	815.2	53.0	872	5	BU521359
42	813	52.9	908	5	BQ953445
43	808.6	52.6	873	5	BQ955064
44	808.6	52.6	971	5	BX397059
45	802.6	52.2	845	4	BG686012

## ALIGNMENTS

RESULT 1	CR615781	2057 bp	mrna	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CSODK012YE23 of Homo sapiens (human).				
DEFINITION	full-length cDNA clone CSODK012YE23 of Homo sapiens (human).				
ACCESSION	CR615781				
VERSION	CR615781.1	GI:50496588			
KEYWORDS	HTC; CNSLT_cDNA.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 2057)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: secref@genoscope.cns.fr - Web: www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..2057				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSODK012YE23"				
	/tissue_type="Hela cells Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	99.3%	Score 1526.8;	DB 3;	Length 2057;	
Best Local Similarity	99.9%	Pred. No. 0;			
Matches 1528;	Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;

Qy	6	GCCCGCGGCGCTCGACGCGACGAGCGAGCGCGAGATGGCGGAGCTTCAGCGCGCG	65
Db	7	GCCCGCGGCGCTCGACGCGACGAGCGCGAGATGGCGGAGCTTCAGCGCGCG	66

66 ATCTCAGAGCTGGTCTGGCTTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 125  
 67 ATCTCAGAGCTGGTCTGGCTTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 126  
 126 GCTCGCTAAGGCTGGCTGGCTTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 185  
 127 GCTCGCTAAGGCTGGCTGGCTTAAGCAAGATTGGAGCTACTCTGCTGTCACCTCCA 186  
 186 GATCCCCCTGCTGGGCGCATGAGCAGCTGCTCTAGAGATTGCTAACTGAGGAAC 245  
 187 GATCCCCCTGCTGGGCGCATGAGCAGCTGCTCTAGAGATTGCTAACTGAGGAAC 246  
 246 CACCAGAGACTCTCAAAATACCACTTTGATGCTCTATGCGGAGAGGAGTACTCAGCAG 305  
 247 CACCAGAGACTCTCAAAATACCACTTTGATGCTCTATGCGGAGAGGAGTACTCAGCAG 306  
 306 GACATCTATCAGGTTTCAGTGCAGCCCATCTAAGGCATTTGCTGGAAGGGCAGATGCC 365  
 307 GACATCTATCAGGTTTCAGTGCAGCCCATCTAAGGCATTTGCTGGAAGGGCAGATGCC 366  
 366 AGTGTGCTTCCCTATGACCCACAGAGCTGGGAAGCGCACAAATGCTGGGCGAGCCCA 425  
 367 AGTGTGCTTCCCTATGACCCACAGAGCTGGGAAGCGCACAAATGCTGGGCGAGCCCA 426  
 426 GAGCAACTCTGGGTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAG 485  
 427 GAGCAACTCTGGGTGATCCCGCGGCTCTCATGACCTCTGAGCTCACAAGGGAGGAG 486  
 486 GGTGCGAGGCGCGGCATCGGCTCTCTCCAGAGCCCATCAGTACTGCTGATTTT 665  
 607 CGGGGAGATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTACTGCTGATTTT 666  
 666 GAGGAGCTATGACCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAGACTGC 605  
 547 GAGAAGTATGACCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAGACTGC 606  
 606 CGGGGAGATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTACTGCTGATTTT 665  
 607 CGGGGAGATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTACTGCTGATTTT 666  
 666 GAGGAGCTATGACCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAGACTGC 725  
 667 GAGGAGCTATGACCTCTGAGCCCTGCTTGGGAGACCTGTAATCCGAGAGACTGC 726  
 726 CGCTCTCCGAGTCACTGCTGCTCCTGCTCAAGGTGACAGCGGAGACCTTTGGCC 785  
 727 CGCTCTCCGAGTCACTGCTGCTCCTGCTCAAGGTGACAGCGGAGACCTTTGGCC 786  
 786 CCATTTCCGAGGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTCCAGAGGACAAC 845  
 787 CCATTTCCGAGGAGAGGAGAACTCTACCTGATTTGACTTGGCTGGGTCCAGAGGACAAC 846  
 846 CGGCGCAGAGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTG 905  
 847 CGGCGCAGAGCAACAGGCGCTTCGGCTAAAGAGAGTGGAGCCATCAACCTCCCTG 906  
 906 TTTGCTCTGGCAAGTGTGAGTGGCTGCTCAAGTGGAGCGGAGACCTTTGGCC 965  
 907 TTTGCTCTGGCAAGTGTGAGTGGCTGCTCAAGTGGAGCGGAGACCTTTGGCC 966  
 966 GACAGCAAGCTCACTCGCTTATGAGGAGTCTCTGCTGGGTGGGTCCAGCAGTATCCTT 1025  
 967 GACAGCAAGCTCACTCGCTTATGAGGAGTCTCTGCTGGGTGGGTCCAGCAGTATCCTT 1026  
 1026 ATTGCCAATATGCGCTGAGAGACGCTTACTAGACACAGTCTCCGACCTCAACTTT 1085  
 1027 ATTGCCAATATGCGCTGAGAGACGCTTACTAGACACAGTCTCCGACCTCAACTTT 1086  
 1086 GCTGCCAGGTCGAAGAGGATGATCAATCGGCTTTTACCAATGAGAGCTGCGAGCTCAT 1145  
 1087 GCTGCCAGGTCGAAGAGGATGATCAATCGGCTTTTACCAATGAGAGCTGCGAGCTCAT 1146

1146 GCCTTGGGAGCTGTTAAGCTGCTCAGAAAGAAATTGCTGTCACCCAGAGGCAAGAGA 1205  
 1147 GCCTTGGGAGCTGTTAAGCTGCTCAGAAAGAAATTGCTGTCACCCAGAGGCAAGAGA 1206  
 1206 GCCCGAGGCGCTGAGAAAGAGGAGATTGGAGCCCTTGGAGCCCATGCGAGCTCTCAGCTCT 1265  
 1207 GCCCGAGGCGCTGAGAAAGAGGAGATTGGAGCCCTTGGAGCCCATGCGAGCTCTCAGCTCT 1266  
 1266 GCCTCCAGAACTCAGCCCTTACAGAACTAAGCAGCATGAGCCCGGCATGCTGGAG 1325  
 1267 GCCTCCAGAACTCAGCCCTTACAGAACTAAGCAGCATGAGCCCGGCATGCTGGAG 1326  
 1326 CGCTCTCAGCTTGGAGCCCTTGGAGCCCTTGGAGCCCATGCGAGCTCTCAGCTCT 1385  
 1327 CGCTCTCAGCTTGGAGCCCTTGGAGCCCTTGGAGCCCATGCGAGCTCTCAGCTCT 1386  
 1386 AGTACCCCAAGCAGAGCGGATGGTGTAAATGAAGACAGTGAAGAGAGGAGCTAGAG 1445  
 1387 AGTACCCCAAGCAGAGCGGATGGTGTAAATGAAGACAGTGAAGAGAGGAGCTAGAG 1446  
 1446 ATTGAGAGCTTAAAGCAGAAAGAACTGGAGCCCAAGATGTTGCCCGAGAGGCT 1505  
 1447 ATTGAGAGCTTAAAGCAGAAAGAACTGGAGCCCAAGATGTTGCCCGAGAGGCT 1506  
 1506 GAGGAAAGAGAGCACTTGTCCCAATG 1535  
 1507 GAGGAAAGAGAGCACTTGTCCCAATG 1536

RESULT 2  
 CR590974  
 LOCUS  
 DEFINITION full-length cDNA clone CS0D1009YJ01 of Placenta Cot 25-normalized of Homo sapiens (human).  
 ACCESSION CR590974  
 VERSION 1  
 KEYWORDS HTc; CNSLT\_cDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2025)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalized  
 JOURNAL Unpublished  
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue  
 REFERENCE 2 (bases 1 to 2025)  
 Genoscope.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
 COMMENT - Web : www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
 FEATURES  
 Location/Qualifiers  
 1..2025  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1009YJ01"  
 /tissue\_type="Placenta Cot 25-normalized"  
 /plasmid="pCMVSPORT\_6"

ORIGIN  
 Query Match 98.4%; Score 1512.8; DB 3; Length 2025;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



```
QY 20 GACGAGCAGAGGCGACCGAGATGGCGGAGCTTACGCGGCGGCGATCTCAGAGCTGG 79
Db 1 GACGAGCAGAGGCGACCGAGATGGCGGAGCTTACGCGGCGGCGATCTCAGAGCTGG 60
QY 80 TCGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCTCAGCTCCAGCTCCGCTAAGGCT 139
Db 61 TCGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCTCAGCTCCAGCTCCGCTAAGGCT 120
QY 140 GGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCTCAGCTCCAGCTCCGCTAAGGCT 199
Db 121 GGCTGTGGCTAAGCAAGATTGGAGCTACTCGTGTCTCAGCTCCAGCTCCGCTAAGGCT 180
QY 200 CGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACCAACCAAGGAGACTCT 259
Db 181 CGGGGCGATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGAACCAACCAAGGAGACTCT 240
QY 260 CAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACTCTATGCGAG 319
Db 241 CAATACCAAGTTGATGCTTCTATGGGAGAGAGTACTCAGCAGGACTCTATGCGAG 300
QY 320 TTCAAGTGCAGCCCATCTTAAGGCACTTCTGGAAGGGCAGAAATGCGAGTGTCTTGCTTA 379
Db 301 TTCAAGTGCAGCCCATCTTAAGGCACTTCTGGAAGGGCAGAAATGCGAGTGTCTTGCTTA 360
QY 380 TGGACCCACAGAGCTGGGAAGACGCAACAATGCTGGGCGAGCCAGAGCAACCTGGGGT 439
Db 361 TGGACCCACAGAGCTGGGAAGACGCAACAATGCTGGGCGAGCCAGAGCAACCTGGGGT 420
QY 440 GATCCGGGGCTCTCATGGAACCTCTGCAAGCTCACAAGGAGAGGCTGCGAGGGCG 499
Db 421 GATCCGGGGCTCTCATGGAACCTCTGCAAGCTCACAAGGAGAGGCTGCGAGGGCG 480
QY 500 GCCATGGGCCCTTCTGTACCATGCTTACTAGAGATCTACAGGAGAGGATATTAGA 559
Db 481 GCCATGGGCCCTTCTGTACCATGCTTACTAGAGATCTACAGGAGAGGATATTAGA 540
QY 560 CCTCTGAGACCTCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGGAATATCCT 619
Db 541 CCTCTGAGACCTCTTGGGAGACCTGGTAATCCGAGAGACTGCCGGGGGAATATCCT 600
QY 620 GATTCGGGGTCTCCGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTCCT 679
Db 601 GATTCGGGGTCTCCGAGAGCCCATCAGTAGCTTTGCTGATTTGAGCGGCACTTCCT 660
QY 680 GCCAGCAGTCGAATCGGACTGTAGAGCACCGGCTCAACAGGCTCCTCCCGGAG 739
Db 661 GCCAGCAGTCGAATCGGACTGTAGAGCACCGGCTCAACAGGCTCCTCCCGGAG 720
QY 740 TCATGCTGTGCTCTGCTCAAGGTGAGCGACGCGGGAAGCTTTGGCCCCCATTTGCCAGCG 799
Db 721 TCATGCTGTGCTCTGCTCAAGGTGAGCGACGCGGGAAGCTTTGGCCCCCATTTGCCAGCG 780
QY 800 AGAGGGAACCTACTGATGATCTTGCTGGGTGAGAGCAACCGGCGCAAGGCAA 859
Db 781 AGAGGGAACCTACTGATGATCTTGCTGGGTGAGAGCAACCGGCGCAAGGCAA 840
QY 860 CAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTCGGCAA 919
Db 841 CAAGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTGCTCGGCAA 900
QY 920 AGTGTAGATGCGTGAATCAGGGCCCTCCCTCGGTATCTTATCGGACAGCAAGCTCAC 979
Db 901 AGTGTAGATGCGTGAATCAGGGCCCTCCCTCGGTATCTTATCGGACAGCAAGCTCAC 960
QY 980 TCGCTTATGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTATGCGCAACATGC 1039
Db 961 TCGCTTATGAGAGCTCTCTGGGTGGCTCAGGCCACAGTATCTTATGCGCAACATGC 1020
QY 1040 CCCTGAGAGAGCTTCTACCTAGACAGCTCTCGCACTCAACTTCTGCTCCAGGTCCTAA 1099
Db 1021 CCCTGAGAGAGCTTCTACCTAGACAGCTCTCGCACTCAACTTCTGCTCCAGGTCCTAA 1080
QY 1100 GGAGGTGATCAATCGGCGCTTTTACCAATGAGAGCTCGAGCCTCATGCTTGGGACCTGT 1159
```

```
Db 1081 GGAGGTGATCAATCGGCGCTTTTACCAATGAGAGCTCGAGCCTCATGCTTTGGGACCTGT 1140
QY 1160 TAAGCTGTCTCAGAAAGAAATTTGCTTGTCTCCACAGAGCAAGAGAGCCCGAGCCCTGA 1219
Db 1141 TAAGCTGTCTCAGAAAGAAATTTGCTTGTCTCCACAGAGCAAGAGAGCCCGAGCCCTGA 1200
QY 1220 GGAAGAGAGATTTGGGAGCCCTGAGCCCATCGGAGCTCCAGCCTCTGCTCTCCAGAAACT 1279
Db 1201 GGAAGAGAGATTTGGGAGCCCTGAGCCCATCGGAGCTCCAGCCTCTGCTCTCCAGAAACT 1260
QY 1280 CAGCCCCCTACAGAAAGCTAAGCAGCATGGACCCCGCCATGCTGGAGCGCCTCTCAGCTT 1339
Db 1261 CAGCCCCCTACAGAAAGCTAAGCAGCATGGACCCCGCCATGCTGGAGCGCCTCTCAGCTT 1320
QY 1340 GGACCGTCTGCTTGCCTCCAGGGGAGCCAGGGGGCCCTCTGTTAGTAGTACCCCAAGCG 1399
Db 1321 GGACCGTCTGCTTGCCTCCAGGGGAGCCAGGGGGCCCTCTGTTAGTAGTACCCCAAGCG 1380
QY 1400 AGACCGATGCTGCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTTAA 1459
Db 1381 AGACCGATGCTGCTTAATGAAGACAGTAGAAGAGAGGACCTAGAGATTGAGAGGCTTAA 1440
QY 1460 GACGAACCAAAAGAACTGGAGGCCAAAGATGTTGCCCCAGAGGCTGAGGAAAAGGAGAA 1519
Db 1441 GACGAACCAAAAGAACTGGAGGCCAAAGATGTTGCCCCAGAGGCTGAGGAAAAGGAGAA 1500
QY 1520 CCATTGTCACCAATG 1535
Db 1501 CCATTGTCACCAATG 1516

RESULT 3
CR590954
LOCUS CR590954 2016 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODI040V106 of Placentia Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR590954
VERSION CR590954.1 GI:50471761
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2016)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 2016)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1. .2016
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI040V106"
/tissue_type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 95.1%; Score 1462.8; DB 3; Length 2016;
```



```

PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2055)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashi, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imokani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kohji, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saioh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
source
1..2055
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D230020105"
/db_xref="taxon:10090"
/clone="D230020105"
/tissue_type="eyeball"
/clone_libs="RIKEN full-length enriched mouse cDNA library"
/dev_stages="12 days embryo"
misc_feature
8..1991
/note="kinesin superfamily protein KIF22
putative"
ORIGIN
Query Match 69.9%; Score 1074.8; DB 3; Length 2055;
Best Local Similarity 83.6%; Pred. No. 1.8e-273;
Matches 1214; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
Qy 59 GCGCGGATCTCAGAGCTGCTGCTGCGCTGAGCAAGATTGAGACTACTCGTCGTC 118
Db 49 GCGCTCGGCCACCTCAGAGCTGCGCGCTGTGTGTAAGCAAGAGGGCTTGGCGCGTC 108
Qy 119 ACCTCCAGCTCGCTGAGGGTGGCTGTGCGACTCGGCCATTCTGTGGATGGAACAGCGGG 178
Db 109 GCCTCTAGCTCGAGTAGGGTGGCTGTGCGCGCTGCGCGCTGTATGATGAGAGACAGA 168
Qy 179 AGCAAGTGAATCCCGCTGTGTGCGGGGCATGGACAGCTGCTCTCTAGAGATTGCTAACTG 238
Db 169 AGCGAAGGAGTCCCGCTGTGTGCGAGCCATAGACAGCTGCTCTCTTGAAGTGGCTACTG 228
Qy 239 GAGGAACACACAGGAGACTCTCAATACCAAGTTGATGCTTCTATGGGAGAGAGTAC 298
Db 229 GAAGAAATACAGGAGAGCTCAATATACAGTTTGTGCTTCTATGCGGAGAGAGAGCAG 288

```

```

Qy 299 TCAGCAGGACATCTATGCAAGTTTCAGTGCAGCCCATCTTAAGGCATCTGCTGGAAGGCA 358
Db 289 TCAGCAGGAAGTCTATGTAGTTTCAGTACAGCCCATCTTAAGGCATCTGCTGGAAGGCA 348
Qy 359 GAATGCCAGTGTGCTTATGACCCACAGGAGCTGGAGAGCGACACAATGCTGGG 418
Db 349 GAATGCCAGTGTACTTGTCTATGGGCTTACTGGGCGAGGAAGACACACAATGCTGGG 408
Qy 419 CAGCCAGAGCAACTGGGTGATCCCGGGCTCTCATGGACCTCTCGCAGCTCACAG 478
Db 409 CAGCCAGAGCAACTGGAGTATCTTCGGGCTCTCATGGACCTCTCGCAGCTCACAG 468
Qy 479 GGAGAGAGGTGCCCGAGGCGCGCATATGGGCTTCTCTGCACCATGTCTTACTAGAGAT 538
Db 469 GGAAGAGAGTCTGAGGAGCGGCCCATATGGAGCTTCTGTCTATGTCTTATAGAGAT 528
Qy 539 CTACAGAGAGAGTATTTAGACCTCTCTGGACCTCTGCTTCCGGAGACCTGGTAAATCCGAGA 598
Db 529 CTACAGAGAGAGTATTTAGACCTCTTGGATCTCTGCATCAGGAGACCTCGTGATCCGGA 588
Qy 599 AGACTGCGGGGGAATATCTGATTCGGGTCTCTCCAGAGGCCCATCAGTAGCTAGCTTTCG 658
Db 589 AGACTGCGGGGGAATATCTGATTCGGGTCTCTCCAGAGGCCCATCAGTAGCTTTCG 648
Qy 659 TGATTTTTCAGGCGCATCTCTCCAGGCGAGTCGAAATCGGACTGTAGGAGCCACCCGGCT 718
Db 649 TGACTTTCAGGAGAGTCTCTCTCCAGGCGAGTCGAAATCGGACTGTAGGAGCCACCCGGCT 708
Qy 719 CAACAGCGCTCTCCCGCAGTCACTGTCTCTCTGCTCAAGTGGACCGGGAACG 778
Db 709 TAACAGCGCTCTCTCCCGTAGTCACGAGTCTCTTGTGTCAAGGTAGATCAGCGTGAACG 768
Qy 779 TTTGGCCCCATTTCCGCGAGGAGGGAATACTCTACTGATTGACTTGGCTGGGTCTAGA 838
Db 769 TTTGACTCCATTTCCGCGAGGGAAGGAACTCTACTTATTGATTGGGTGGTCTAGA 828
Qy 839 GGACAAACCGGCGCAGGCAACAGGGCTTCGGCTAAAGAGAGTGAGGCCATCAACAC 898
Db 829 GGACAAACCGTCCGACAGGCAACAGGGCATTCGGCTAAAGAGAGCGGAGCCATCAACAC 888
Qy 899 CTCCTCTTTGCTCTCGGCAAAAGTGTAGATGCGCTGAATCAGGCGCTCTCTCGTGTACC 958
Db 889 CTCCTCTTTGCTCTCGGCAAAAGTGTAGATGCGCTGAATCAGGCGCTCTCTCGCATACC 948
Qy 959 TTATCGGAGCAGCAGCTCACTCGCCTATTGCGAGACTCTCTGGGTGCTCAGCCACAG 1018
Db 949 ATACCGGAGCAGCAGCTCACTCGCTCTGCTGCGAGACTCTCTGGGAGGCTCAGCTCATAG 1008
Qy 1019 TATCCTTATTGCGCAACATTGCGCTGAGAGAGCTTCTTACCTAGACAGTCTCCGCACT 1078
Db 1009 CATCTCTATTGCGCAACATTGCTCTGAAAGAGCTTTTACAGGATACATTTTCAGCATT 1068
Qy 1079 CAACCTTTGCTCGCAGGTCCAAGAGGTGATCAATCGGCTTTTACCAATGAGAGCTGCA 1138
Db 1069 GAACCTTCACTGCTAGGTCCAAGAGGTGATTAACCGGCTTTTACCAATGAGAGTCTACA 1128
Qy 1139 GCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCACAGAGGC 1198
Db 1129 GCCTCATGCTTGGGACCTGTTAAGCTGTCTCAGAAAGAAATGCTTGGTCCGCTCAGAGGC 1188
Qy 1199 AAAGAGAGCCCGGCGCTGAGGAAGAGGATTTGGGAGCCCTGAGCCCATGGCAGCTCC 1258
Db 1189 AAAGAGAGCCCGGCGCTGAGGAAGAAATCACTGGGAGCCCTGAGTCTACAGAGCTCC 1248
Qy 1259 AGCTCTGCTCTCCAGAAACTCAGCCCTTACAGAACTAAGCAGCATGGACCCGGCCAT 1318
Db 1249 AGCTCTGCTCTCCAGAAACTAAGCTCTTACAGAGCTAAGCAATATGGATCCAGCCAT 1308
Qy 1319 GCTGAGAGCGCTCTCTAGCTTGGACCGTCTGCTTCCCTCCAGGGAGCCAGGGGGCCCC 1378
Db 1309 GTTGGAGAACTCTCTGAGTATGGAACGTTTGTGGTTCCTCCAGGGAAGCCAGGAGCCCC 1368

```



Qy 599 AGACTGCCGGGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 658  
 Db 592 AGACTGCCGGGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 651  
 Qy 659 TGATTTTGGGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 718  
 Db 652 TGATTTTGGGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 711  
 Qy 719 CAACAGGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 778  
 Db 712 TAACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 771  
 Qy 779 TTTGGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 838  
 Db 772 TTTGGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 831  
 Qy 839 GGACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 898  
 Db 832 GGACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 891  
 Qy 899 CTCCTGTTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 958  
 Db 892 CTCCTGTTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 951  
 Qy 959 TTATCGGACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1018  
 Db 952 TTATCGGACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1011  
 Qy 1019 TATCGGACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1078  
 Db 1012 TATCGGACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1071  
 Qy 1079 CAACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1138  
 Db 1072 CAACAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1131  
 Qy 1139 GCCTCATGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1198  
 Db 1132 GCCTCATGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1191  
 Qy 1199 AAAGAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1258  
 Db 1192 AAAGAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1251  
 Qy 1259 AGCTCTGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1318  
 Db 1252 AGCTCTGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1311  
 Qy 1319 GCTGAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1378  
 Db 1312 GCTGAGGCTCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1371  
 Qy 1379 TCTGTTGATTTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1438  
 Db 1372 TCTGTTGATTTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1431  
 Qy 1439 CTTAGAGTTGAGGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1498  
 Db 1432 CTTAGAGTTGAGGCTCTCCGGGATATCTCTGATTTCCGGGCTCTCCAGAGGCCATCAGTGTTC 1491  
 Qy 1499 GAAGGCTGAGGA 1510  
 Db 1492 GGAGGCTGAGGA 1503

RESULT 6  
 AL545712  
 LOCUS  
 DEFINITION AL545712 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
 clone CS001009Y001 5-PRIME, mRNA sequence.  
 AL545712  
 AL545712.3 GI:45746191

KEYWORDS  
 SOURCE  
 ORGANISM

Homo sapiens (human)  
 Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1124)  
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)

On Feb 15, 2001 this sequence version replaced gi:31267547.

Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with NotI and cloned  
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 7766.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?S=CS001009CE01Q1&c=7766.r.

FEATURES  
 source

Location/Qualifiers  
 1..1124  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS001009Y001"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 7766.r"

ORIGIN

Query Match 68.6%; Score 1054.4; DB 1; Length 1124;  
 Best Local Similarity 96.5%; Pred. No. 3.9e-268;  
 Matches 1068; Conservative 25; Mismatches 12; Indels 2; Gaps 2;  
 Qy 20 GAGCAGCAGAGGCGAGGATGGCGGAGCTTACGCGCGCGCGAGTCTCAGAGCTGG 79  
 Db 1 GAGCAGCAGAGGCGAGGATGGCGGAGCTTACGCGCGCGCGAGTCTCAGAGCTGG 60  
 Qy 80 TCGTGTGCGTAAAGAGATTTGGAGCTACTCGTGTCCACCTCCAGCTCGCTAAAGGT 139  
 Db 61 TCGTGTGCGTAAAGAGATTTGGAGCTACTCGTGTCCACCTCCAGCTCGCTAAAGGT 120  
 Qy 140 GGCTGTGCGTAAAGAGATTTGGAGCTACTCGTGTCCACCTCCAGCTCGCTAAAGGT 199  
 Db 121 GGCTGTGCGTAAAGAGATTTGGAGCTACTCGTGTCCACCTCCAGCTCGCTAAAGGT 180  
 Qy 200 GCGGGGATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAAACCAAGAGAGCTCT 259  
 Db 181 GCGGGGATGGACAGCTGCTCTTAGAGATTGCTAACTGGAGGAAACCAAGAGAGCTCT 240  
 Qy 260 CAATACACGTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGAGG 319  
 Db 241 CAATACACGTTGATGCTCTTATGGGAGAGAGTACTCAGCAGGACATCTATGAGG 300  
 Qy 320 TTCACTGACGCGCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTT 379  
 Db 301 TTCACTGACGCGCCATCTTAAGGCACTTGTGGAAGGGCAGAAATGCCAGTGTGCTT 360  
 Qy 380 TGGACCCACAGGAGCTGGGAAGAGCGACACAATGCTGGGAGGCCAGAGCAACCTGGGT 439  
 Db 361 TGGACCCACAGGAGCTGGGAAGAGCGACACAATGCTGGGAGGCCAGAGCAACCTGGGT 420  
 Qy 440 GATCCCGCGGCTCTCATGACCTCTCTGAGCTCTCAAGGGAGAGGAGTCCCGAGGCCG 499  
 Db 421 GATCCCGCGGCTCTCATGACCTCTCTGAGCTCTCAAGGGAGAGGAGTCCCGAGGCCG 480  
 Qy 500 GCCATGGGCGCTTCTCTCACCAGTCTTACCTAGAGATCTACCCAGGAGAGGATTAGA 559

```

481  GCCATGGCCCTTCTCTGCTACCATGCTTACCTAGAGATCTACGAGAGAGGKATAGA 540
482  |||
483  |||
484  |||
485  |||
486  |||
487  |||
488  |||
489  |||
490  |||
491  |||
492  |||
493  |||
494  |||
495  |||
496  |||
497  |||
498  |||
499  |||
500  |||
501  |||
502  |||
503  |||
504  |||
505  |||
506  |||
507  |||
508  |||
509  |||
510  |||
511  |||
512  |||
513  |||
514  |||
515  |||
516  |||
517  |||
518  |||
519  |||
520  |||
521  |||
522  |||
523  |||
524  |||
525  |||
526  |||
527  |||
528  |||
529  |||
530  |||
531  |||
532  |||
533  |||
534  |||
535  |||
536  |||
537  |||
538  |||
539  |||
540  |||
541  |||
542  |||
543  |||
544  |||
545  |||
546  |||
547  |||
548  |||
549  |||
550  |||
551  |||
552  |||
553  |||
554  |||
555  |||
556  |||
557  |||
558  |||
559  |||
560  |||
561  |||
562  |||
563  |||
564  |||
565  |||
566  |||
567  |||
568  |||
569  |||
570  |||
571  |||
572  |||
573  |||
574  |||
575  |||
576  |||
577  |||
578  |||
579  |||
580  |||
581  |||
582  |||
583  |||
584  |||
585  |||
586  |||
587  |||
588  |||
589  |||
590  |||
591  |||
592  |||
593  |||
594  |||
595  |||
596  |||
597  |||
598  |||
599  |||
600  |||
601  |||
602  |||
603  |||
604  |||
605  |||
606  |||
607  |||
608  |||
609  |||
610  |||
611  |||
612  |||
613  |||
614  |||
615  |||
616  |||
617  |||
618  |||
619  |||
620  |||
621  |||
622  |||
623  |||
624  |||
625  |||
626  |||
627  |||
628  |||
629  |||
630  |||
631  |||
632  |||
633  |||
634  |||
635  |||
636  |||
637  |||
638  |||
639  |||
640  |||
641  |||
642  |||
643  |||
644  |||
645  |||
646  |||
647  |||
648  |||
649  |||
650  |||
651  |||
652  |||
653  |||
654  |||
655  |||
656  |||
657  |||
658  |||
659  |||
660  |||
661  |||
662  |||
663  |||
664  |||
665  |||
666  |||
667  |||
668  |||
669  |||
670  |||
671  |||
672  |||
673  |||
674  |||
675  |||
676  |||
677  |||
678  |||
679  |||
680  |||
681  |||
682  |||
683  |||
684  |||
685  |||
686  |||
687  |||
688  |||
689  |||
690  |||
691  |||
692  |||
693  |||
694  |||
695  |||
696  |||
697  |||
698  |||
699  |||
700  |||
701  |||
702  |||
703  |||
704  |||
705  |||
706  |||
707  |||
708  |||
709  |||
710  |||
711  |||
712  |||
713  |||
714  |||
715  |||
716  |||
717  |||
718  |||
719  |||
720  |||
721  |||
722  |||
723  |||
724  |||
725  |||
726  |||
727  |||
728  |||
729  |||
730  |||
731  |||
732  |||
733  |||
734  |||
735  |||
736  |||
737  |||
738  |||
739  |||
740  |||
741  |||
742  |||
743  |||
744  |||
745  |||
746  |||
747  |||
748  |||
749  |||
750  |||
751  |||
752  |||
753  |||
754  |||
755  |||
756  |||
757  |||
758  |||
759  |||
760  |||
761  |||
762  |||
763  |||
764  |||
765  |||
766  |||
767  |||
768  |||
769  |||
770  |||
771  |||
772  |||
773  |||
774  |||
775  |||
776  |||
777  |||
778  |||
779  |||
780  |||
781  |||
782  |||
783  |||
784  |||
785  |||
786  |||
787  |||
788  |||
789  |||
790  |||
791  |||
792  |||
793  |||
794  |||
795  |||
796  |||
797  |||
798  |||
799  |||
800  |||
801  |||
802  |||
803  |||
804  |||
805  |||
806  |||
807  |||
808  |||
809  |||
810  |||
811  |||
812  |||
813  |||
814  |||
815  |||
816  |||
817  |||
818  |||
819  |||
820  |||
821  |||
822  |||
823  |||
824  |||
825  |||
826  |||
827  |||
828  |||
829  |||
830  |||
831  |||
832  |||
833  |||
834  |||
835  |||
836  |||
837  |||
838  |||
839  |||
840  |||
841  |||
842  |||
843  |||
844  |||
845  |||
846  |||
847  |||
848  |||
849  |||
850  |||
851  |||
852  |||
853  |||
854  |||
855  |||
856  |||
857  |||
858  |||
859  |||
860  |||
861  |||
862  |||
863  |||
864  |||
865  |||
866  |||
867  |||
868  |||
869  |||
870  |||
871  |||
872  |||
873  |||
874  |||
875  |||
876  |||
877  |||
878  |||
879  |||
880  |||
881  |||
882  |||
883  |||
884  |||
885  |||
886  |||
887  |||
888  |||
889  |||
890  |||
891  |||
892  |||
893  |||
894  |||
895  |||
896  |||
897  |||
898  |||
899  |||
900  |||
901  |||
902  |||
903  |||
904  |||
905  |||
906  |||
907  |||
908  |||
909  |||
910  |||
911  |||
912  |||
913  |||
914  |||
915  |||
916  |||
917  |||
918  |||
919  |||
920  |||
921  |||
922  |||
923  |||
924  |||
925  |||
926  |||
927  |||
928  |||
929  |||
930  |||
931  |||
932  |||
933  |||
934  |||
935  |||
936  |||
937  |||
938  |||
939  |||
940  |||
941  |||
942  |||
943  |||
944  |||
945  |||
946  |||
947  |||
948  |||
949  |||
950  |||
951  |||
952  |||
953  |||
954  |||
955  |||
956  |||
957  |||
958  |||
959  |||
960  |||
961  |||
962  |||
963  |||
964  |||
965  |||
966  |||
967  |||
968  |||
969  |||
970  |||
971  |||
972  |||
973  |||
974  |||
975  |||
976  |||
977  |||
978  |||
979  |||
980  |||
981  |||
982  |||
983  |||
984  |||
985  |||
986  |||
987  |||
988  |||
989  |||
990  |||
991  |||
992  |||
993  |||
994  |||
995  |||
996  |||
997  |||
998  |||
999  |||
1000  |||

```

```

RESULT 7
BX337406
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

BX337406 1027 bp mRNA linear EST 07-APR-2004  
 BX337406 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0DI041V113 5-PRIME, mRNA sequence.  
 BX337406.2 GI:46274820  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1027)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 2, 2003 this sequence version replaced gi:30341614.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 7766.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0DI041AB07QP1&c=7766.r.

```

FEATURES
source
Location/Qualifiers
1..1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI041V113"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 65.2%; Score 1002.8; DB 5; Length 1027;
Best Local Similarity 99.5%; Pred. No. 1.8e-254;
Matches 1023; Conservative 3; Mismatches 0; Indels 2; Gaps 2;
QY 38 CGAGATGGCGGAGCTTTCAGC-GGGCGGATCTCAGGAGCTGCTCGCTCGGCTAAGCA 96
DB 1 CGAGATGGCGGAGCTTTCAGC-GGGCGGATCTCAGGAGCTGCTCGCTCGGCTAAGCA 60
QY 97 AGATTGGAGCTACTCGTCTGCTCCAGCTCCAGCTCGGCTGAGGGTGGCTGTCGACTCGGC 156
DB 61 AGATTGGAGCTACTCGTCTGCTCCAGCTCCAGCTCGGCTGAGGGTGGCTGTCGACTCGGC 120
QY 157 CATTTGGATGGAACAGCGGAGCAAGTGAATCCCTCTGTTGTCGGGGATGGAACAGCT 216
DB 121 CATTTGGATGGAACAGCGGAGCAAGTGAATCCCTCTGTTGTCGGGGATGGAACAGCT 180
QY 217 GCTCTCTAGAGATGCTTAACTGGAGAACACAGGAGACTCTCAAAATACCAAGTTGATG 276
DB 181 GCTCTCTAGAGATGCTTAACTGGAGAACACAGGAGACTCTCAAAATACCAAGTTGATG 240
QY 277 CTTTCTATGGGAGAGAGTACTCAGAGGAGATCTATCAGAGTTTCAGTGCAGCCATCC 336
DB 241 CTTTCTATGGGAGAGAGTACTCAGAGGAGATCTATCAGAGTTTCAGTGCAGCCATCC 300
QY 337 TAAGGACATTGCTGGAAGGGCAGATGCGAGTGTGCTTGCCTATGACCCACAGAGAGCTG 396
DB 301 TAAGGACATTGCTGGAAGGGCAGATGCGAGTGTGCTTGCCTATGACCCACAGAGAGCTG 360
QY 397 GGAAGACGACACAAATGCTGGGAGAGAGAGTGTGCGGGTGTATCCCGGGGCTCTCA 456
DB 361 GGAAGACGACACAAATGCTGGGAGAGAGAGTGTGCGGGTGTATCCCGGGGCTCTCA 420
QY 457 TGGACCTCTGCACTCACAAGGAGAGAGTGTGCGGGGAGAGTGTGCGGGGCTCTCTG 516
DB 421 TGGACCTCTGCACTCACAAGGAGAGAGTGTGCGGGGAGAGTGTGCGGGGCTCTCTG 479
QY 517 TCACCATGTTTACCTAGAGATCTACAGGAGAGAGTATTAGACCTCTGACCTGCTT 576
DB 480 TCACCATGTTTACCTAGAGATCTACAGGAGAGAGTATTAGACCTCTGACCTGCTT 539
QY 577 CGGAGACCTGTTAATCCGAGAGAGTGTGCGGGGAGAGTATCTCTGATTCGCGGTCTCCC 636
DB 540 CGGAGACCTGTTAATCCGAGAGAGTGTGCGGGGAGAGTATCTCTGATTCGCGGTCTCCC 599
QY 637 AGAAGCCCATCATGAGCTTTGCTGATTTTGAAGCGGACCTCTCTGCGAGCAGTGAATC 696
DB 600 AGAAGCCCATCATGAGCTTTGCTGATTTTGAAGCGGACCTCTCTGCGAGCAGTGAATC 659
QY 697 GGAATGTAGAGAGAGAGAGTGTGCGGGGAGAGTGTGCGGGGAGAGTGTGCGGTCTG 756
DB 660 GGAATGTAGAGAGAGAGAGTGTGCGGGGAGAGTGTGCGGGGAGAGTGTGCGGTCTG 719
QY 757 TCAAGGTGAGAGAGAGAGTGTGCGGGGAGAGTGTGCGGGGAGAGTGTGCGGTCTG 816
DB 720 TCAAGGTGAGAGAGAGAGTGTGCGGGGAGAGTGTGCGGGGAGAGTGTGCGGTCTG 779
QY 817 TGATTGACTTGGCTGGGTGAGAGGACAAACCGGCGCAGAGGCAACAGGCGCTTGGCGTAA 876
DB 780 TGATTGACTTGGCTGGGTGAGAGGACAAACCGGCGCAGAGGCAACAGGCGCTTGGCGTAA 839

```

```

Qy 877 AAGAGGTGGAGCCATCAACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATGGCTGA 936
Db 840 AAGAGGTGGAGCCATCAACCTCCCTGTTTGTCTGGGCAAAAGTGTAGATGGCTGA 899
Qy 937 ATCAGGCGCTCCCTGTTAGCTTATCGGACAGCAAGCTCACTGCGCTATTGAGGACT 996
Db 900 ATCAGGCGCTCCCTGTTAGCTTATCGGACAGCAAGCTCACTGCGCTATTGAGGACT 959
Qy 997 CTCTGGGTGGCTCAGCCACAGTATCTTATTGCAACATTCGCCCTGAGAGAGCTTCT 1056
Db 960 CTCTGGGTGGCTCAGCCACAGTATCTTATTGCAACATTCGCCCTGAGAGAGCTTCT 1019
Qy 1057 ACCTAGAC 1064
Db 1020 ACCTAGAC 1027

RESULT 8
AL559772 1083 bp mRNA linear EST 02-APR-2004
LOCUS clone CS0D005Y11 5-PRIME, mRNA sequence.
DEFINITION AL559772
ACCESSION AL559772.3 GI:46185144
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 1083)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31283903.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7766. r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D005Y11&c=7766.r.
FEATURES
Location/Qualifiers
1..1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D005Y11"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 63.8%; Score 980.8; DB 1; Length 1083;
Best Local Similarity 96.4%; Pred. No. 1.3e-248;
Matches 1010; Conservative 13; Mismatches 21; Indels 4; Gaps 2;

Qy 10 CGGCGCGCTCAGCAGAGGCGAGATGGCGGAGCTTCAGCGCGCGGATCT 69
Db 1 CGGCGCGCTCAGCAGAGGCGAGATGGCGGAGCTTCAGCGCGCGGATCT 60
Qy 70 CAGGAGCTGGTCTGCGGTAAGCAAGATTGGAGTACTCGTCTCCACTCCAGCTC 129

```

```

Db 61 CAGGAGCTGTCGCTGTCGCTAAAGCAAGATTGGAGTACTCTGTCGTCCACCTCCAGCTC 120
Qy 130 GCGTAAGGGTGGCTGTCGACTCGCGGCATTTGTGGATGGAAACGCGGAGCAAGTATC 189
Db 121 GCGTAAGGGTGGCTGTCGACTCGCGGCATTTGTGGATGGAAACGCGGAGCAAGTATC 180
Qy 190 CCCCCTGTGTGCGGGGATGGACAGCTCTCTCTAGAGATTGCTAACTGGAGGAACCAACC 249
Db 181 CCCCCTGTGTGCGGGGATGGACAGCTCTCTCTAGAGATTGCTAACTGGAGGAACCAACC 240
Qy 250 AGGAGACTCTCAAAATACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACA 309
Db 241 AGGAGACTCTCAAAATACAGTTTGTATGCTTCTATGGGAGAGAGTACTCAGCAGGACA 300
Qy 310 TCTATGCAAGTTTCAAGTCAGCCCATCTTAAGGCACTTCTGTAAGGGGAGAGTATGCAAGT 369
Db 301 TCTATGCAAGTTTCAAGTCAGCCCATCTTAAGGCACTTCTGTAAGGGGAGAGTATGCAAGT 360
Qy 370 TGCTTGCTTATGGACCCACAGGAGCTGGGAAGACGACACAATGCTTGGGCAAGCCAGAGC 429
Db 361 TGCTTGCTTATGGACCCACAGGAGCTGGGAAGACGACACAATGCTTGGGCAAGCCAGAGC 420
Qy 430 AACCTGGGGTGTATCCCGGGGCTCTCATGACCTCTCTGAGCTCACAAGGAGAGAGGCTG 489
Db 421 AACCTGGGGTGTATCCCGGGGCTCTCATGACCTCTCTGAGCTCACAAGGAGAGAGGCTG 480
Qy 490 CCGAGGGCGGCCATGGGCCCTTCTGTCACCATGCTTACCTAGAGATCTTACCAGGAGA 549
Db 481 CCGAGGGCGGCCATGGGCCCTTCTGTCACCATGCTTACCTAGAGATCTTACCAGGAGA 540
Qy 550 AGGTATTAGACCTCTCTCGGAGACCTTCTCGGAGACCTGGTAATCCGAGAACTGCCCCGG 609
Db 541 AGGTATTAGACCTCTCTCGGAGACCTTCTCGGAGACCTGGTAATCCGAGAACTGCCCCGG 600
Qy 610 GGAATATCTGATTCGGGCTCTCTCCAGAAAGCCATCAGTAGCTTGTGCTGATTTGAGC 669
Db 601 GGAATATCTGATTCGGGCTCTCTCCAGAAAGCCATCAGTAGCTTGTGCTGATTTGAGC 660
Qy 670 GGCACTTCTCGGAGCAGTCGAAATCGGAGTGGAGCCACCCGGCTCAACAGCGCT 729
Db 661 GGCACTTCTCGGAGCAGTCGAAATCGGAGTGGAGCCACCCGGCTCAAMMAGCGCT 720
Qy 730 CCTCCCGCAGTCATGCTGCTCTCTGTCGAGTGGACCAAGCGGGAACGTTTGGCCCCAT 789
Db 721 CCTCCCGCAGTCATGCTGCTCTCTGTCGAGTGGACCAAGCGGGAACGTTTGGCCCCAT 780
Qy 790 TTCGCCAGCAGAGGGGAAAATCTCTACCTGATTTGCTTGGCTGGCTCAGAGACCAACCGGC 849
Db 781 TTCGCCAGCAGAGGGGAAAATCTCTACCTGATTTGCTTGGCTGGCTCAGAGACCAACCGGC 840
Qy 850 GCACAGGCAACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTG 909
Db 841 GCACAGGCAACAGGGGCTTCGGCTAAAGAGAGTGGAGCCATCAACACCTCCCTGTTTG 900
Qy 910 TCCTGGGCAAAAGTGTAGATGCTGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACA 969
Db 901 TCCTGGGCAAAAGTGTAGATGCTGCTGAATCAGGCGCTCCCTCGTGTACCTTATCGGACA 960
Qy 970 GCAAGCTCACTCGCCTATTTCAGAGACTCTCTGGGTGGCTCAGCCCAAGTATCTTATTG 1029
Db 961 GCAACTCACTCGCCTATTGCA-GRCTCTCTGGGTGGTGCACCCCMAGTAT---CATTATTG 1016
Qy 1030 CCAACATTCGCCCTGAGAGAGCTTCTTA 1057
Db 1017 CAAMATTGCCCTKARAPAGCTTTTCTA 1044

RESULT 9
BX337353 1032 bp mRNA linear EST 07-APR-2004
LOCUS clone CS0D1040Y106 5-PRIME, mRNA sequence.
DEFINITION BX337353
ACCESSION BX337353

```



```
VERSION BX37353.2 GI:46272059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1032)
JOURNAL Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
On May 2, 2003 this sequence version replaced gi:30337600.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1040B030P1&c=7766.r.
FEATURES
Location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1040Y106"
/tissue types="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."
ORIGIN
Query Match 63.6%; Score 978.6; DB 5; Length 1032;
Best Local Similarity 98.6%; Pred. No. 4.8e-248;
Matches 1003; Conservative 4; Mismatches 8; Indels 2; Gaps 2;
QY 70 CAGGAGCTGTCGCTGCTGAGCAAGATGAGAGTCTCTGCTGCTCACTCCAGCTC 129
DB 1 CAGGAGCTGTCGCTGCTGAGCAAGATGAGAGTCTCTGCTGCTCACTCCAGCTC 60
QY 130 GCCTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
DB 61 GCCTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 190 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
DB 121 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 250 AGGAGACTCTCAATACCAATACCAATACCAATACCAATACCAATACCAATACCA 309
DB 181 AGGAGACTCTCAATACCAATACCAATACCAATACCAATACCAATACCAATACCA 240
QY 310 TCTATGAGCTTCAGTGCAGCCATCTCAAGGCACTTCTGAGAGTCTGAGAGTCTGAG 369
DB 241 TCTATGAGCTTCAGTGCAGCCATCTCAAGGCACTTCTGAGAGTCTGAGAGTCTGAG 300
QY 370 TGCTTGCCTATGAGCCACAGGAGCTGGGAGAGCGCACCAATGCTGGGAGCCAGAGC 429
DB 301 TGCTTGCCTATGAGCCACAGGAGCTGGGAGAGCGCACCAATGCTGGGAGCCAGAGC 360
QY 430 AACCTGGGCTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGGAGG 489
DB 361 AACCTGGGCTGATCCCGCGGGCTCTCATGGACCTCTCTGAGCTCAACAAGGAGGAGG 420
QY 490 CCAGAGGCGCGCCATGGGCGCCCTTCTGTCACCAATGCTTACCTAGAGATCTACAGAGA 549
DB 421 CCAGAGGCGCGCCATGGGCGCCCTTCTGTCACCAATGCTTACCTAGAGATCTACAGAGA 480
```

```
550 AGCTATTAGACCTCTCGGACCTCTCGGAGACCTGTAATCCGAGAGACTGCGGG 609
481 AGGTATTAGACCTCTCGGACCTCTCGGAGACCTGTAATCCGAGAGACTGCGGG 540
610 GGAATATCTGATTCGGGCTCTCCAGAGGCCATCAGTAGCTTGTGCTGATTTGAGC 669
541 GGAATATCTGATTCGGGCTCTCCAGAGGCCATCAGTAGCTTGTGCTGATTTGAGC 600
670 GGCATCTCTGCGCAGCAGTCGAAATCGGACTGTAGAGCCACCGGCTCAACAGCGCT 729
601 GGCATCTCTGCGCAGCAGTCGAAATCGGACTGTAGAGCCACCGGCTCAACAGCGCT 660
730 CTCTCCGCGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
661 CTCTCCGCGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
790 TTCTCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
721 TTCTCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
850 GCACAGGCAACAGGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
781 GCACAGGCAACAGGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
910 TCTCTGGGCAAGTGTAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968
841 TCTCTGGGCAAGTGTAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
969 AGCAAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027
901 AGCAAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
1028 TGCAACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
961 TGCAACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017

RESULT 10
BX363358 959 bp mRNA linear EST 08-APR-2004
DEFINITION BX363358 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0D012YE23 5-PRIME, mRNA sequence.
ACCESSION BX363358
VERSION BX363358.2 GI:46287820
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 959)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30368690.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7766.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D012AC12QP1&c=7766.r.
FEATURES
Location/Qualifiers
1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D012YE23"
/cell_type="HELA CELLS COT 25-NORMALIZED"
```



/cell_line="HELA"	
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"	
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
ORIGIN	
Query Match	61.2%; Score 942; DB 5; Length 959;
Best Local Similarity	99.9%; Pred. No. 2.4e-238;
Matches	953; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy	6 GCCCGCGCGGCTCAGCAGAGGCGAGCGCGAGATGGCGGACGCTTCAGCGG-CGCG 64
Db	6 GCCCGCGCGGCTCAGCAGAGGCGAGCGCGAGATGGCGGACGCTTCAGCGGCGCG 65
Qy	65 GATCTCAGAGCTGTGCTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACTCC 124
Db	66 GATCTCAGAGCTGTGCTCGGCTAAGCAAGATTGGAGCTACTGCTGCTCCACTCC 125
Qy	125 AGCTCGCTAAGGGTGGCTGTCGACTCGGCGCAATTTGTGGATGGAACAGCGGAGCAAG 184
Db	126 AGCTCGCTAAGGGTGGCTGTCGACTCGGCGCAATTTGTGGATGGAACAGCGGAGCAAG 185
Qy	185 TGATCCCCCTGTGTGCGGGATGGAACAGCTGCTCTTAGAGATTGCTAACTGGAGAA 244
Db	186 TGATCCCCCTGTGTGCGGGATGGAACAGCTGCTCTTAGAGATTGCTAACTGGAGAA 245
Qy	245 CCACGAGAGACTCTCAATACCACTTTGATGCTTCTATGGGAGAGGAGTACTCAGCA 304
Db	246 CCACGAGAGACTCTCAATACCACTTTGATGCTTCTATGGGAGAGGAGTACTCAGCA 305
Qy	305 GGACATCTATGACGTTCACTGAGCGCCATCTTAAGGCACCTTGTGGAGGCGAGATGC 364
Db	306 GGACATCTATGACGTTCACTGAGCGCCATCTTAAGGCACCTTGTGGAGGCGAGATGC 365
Qy	365 CAGTGTGCTTCCCTATGAGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCC 424
Db	366 CAGTGTGCTTCCCTATGAGACCCACAGGAGCTGGGAAGACGACACAATGCTGGGAGCCC 425
Qy	425 AGACCACTGGGTGATCCCGGGCTCTCATGACCTCTGACGCTCAGAGTCAAGAGGAGGA 484
Db	426 AGACCACTGGGTGATCCCGGGCTCTCATGACCTCTGACGCTCAGAGTCAAGAGGAGGA 485
Qy	485 GGGTCCGAGGCGCGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCA 544
Db	486 GGGTCCGAGGCGCGCCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACCA 545
Qy	545 GGAGAAAGTATTAGACCTCTGGACCCCTGCTTCGGGAGACCTGTTAATCCGAGAGACTG 604
Db	546 GGAGAAAGTATTAGACCTCTGGACCCCTGCTTCGGGAGACCTGTTAATCCGAGAGACTG 605
Qy	605 CCGGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTT 664
Db	606 CCGGGGGAATATCTGATTCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTT 665
Qy	665 TGAGCGGCACCTTCTGCGCAGCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCA 724
Db	666 TGAGCGGCACCTTCTGCGCAGCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCA 725
Qy	725 GCGTCTCTCCGAGTCACTGCTGCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGC 784
Db	726 GCGTCTCTCCGAGTCACTGCTGCTCTGCTCAAGGTGGAACAGCGGGAACGTTTGGC 785
Qy	785 CCATTTTCGCGCAGAGAGGGAACCTCTACTGATTTGAATTGGCTGGGTGAGAGCAAA 844
Db	786 CCATTTTCGCGCAGAGAGGGAACCTCTACTGATTTGAATTGGCTGGGTGAGAGCAAA 845
Qy	845 CCGGCGCACAGGCAACAGGSCCTTCGCTCTAAAGAGGTGGAGCCATCAACCTCCCT 904
Db	846 CCGGCGCACAGGCAACAGGSCCTTCGCTCTAAAGAGGTGGAGCCATCAACCTCCCT 905
Qy	905 GTTTGTCTTGGGCAAGTGGTAGATGGCTGAAATCAGGGCCCTCCCTCGTGATCC 958
Db	906 GTTTGTCTTGGGCAAGTGGTAGATGGCTGAAATCAGGGCCCTCCCTCGTGATCC 959
RESULT 11	
LOCUS	BX354166
DEFINITION	BX354166 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
ACCESSION	BX354166
VERSION	BX354166.2
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On May 5, 2003 this sequence version replaced gi:30373827. Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7766.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0DC015DD02QPl&c=7766.r.
FEATURES	Location/Qualifiers
source	1..958 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DC015VH04" /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED" /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	
Query Match	60.7%; Score 934; DB 5; Length 958;
Best Local Similarity	99.8%; Pred. No. 3.2e-236;
Matches	956; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy	7 CCGCGGCGGCTCGACGACGAGGCGAGCGCGAGATGGCGGACGCTTCAGCGG-CGCG 65
Db	1 CCGCGGCGGCTCGACGACGAGGCGAGCGCGAGATGGCGGACGCTTCAGCGGCGGCG 60
Qy	66 ATCTCAGAGCTGCTCGCTAAGCAAGATTGGAGCTACTGCTGCTCCACTCCA 125
Db	61 ATCTCAGAGCTGCTCGCTAAGCAAGATTGGAGCTACTGCTGCTCCACTCCA 120
Qy	126 GCTCGCTAAGGGTGGCTGTCGACTCCGCGCAATTTGTGGATGGAACAGCGGAGCAAGT 185
Db	121 GCTCGCTAAGGGTGGCTGTCGACTCCGCGCAATTTGTGGATGGAACAGCGGAGCAAGT 180
Qy	186 GATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTAGAGATTGCTAAGTGGAGAAC 245
Db	181 GATCCCCCTGTGTCGGGCGATGGACAGCTGCTCTAGAGATTGCTAAGTGGAGAAC 240
Qy	246 CACGAGAGACTCTCAATACCACTTTGATGCTTCTATGGGAGAGAGTACTCAGCAG 305
Db	241 CACGAGAGACTCTCAATACCACTTTGATGCTTCTATGGGAGAGAGTACTCAGCAG 300
Qy	306 GACATCTATGAGGTTTCAGTGCAGCCCATCTTAAGGCATTTGCTGGAGGCGCAATGCC 365

	301	GACATCTATGCAAGTTTCAAGTCAGCCCATCTTAAGGCACACTTGCTGGAAAGGGCAGAAATGCC	360
Db			
	366	AGTGTGCTTGCTTATGGACCCACACAGGAGCTGGGAAGACGCACACAATGCTGGCGAGCCCCA	425
Qy			
	361	AGTGTGCTTGCTTATGGACCCACACAGGAGCTGGGAAGACGCACACAATGCTGGCGAGCCCCA	420
Db			
	426	GAGCAACTTGGGGTGATCCC CGGGCTCTCATNGACCTCTTCGACCTCACAAAGGAGGAG	485
Qy			
	421	GAGCAAACCTTGGGGTGATCCC CGGGCTCTCATNGACCTCTTCGACCTCACAAAGGAGGAG	480
Db			
	486	GGTCCCGAGGGCCGGCCATGGCCCTTCTCTGCTCAACATGCTTACCTAGAGATCTACCAG	545
Qy			
	481	GGTCCCGAGGGCCGGCCATGGCCCTTCTCTGCTCAACATGCTTACCTAGAGATCTACCAG	540
Db			
	546	GAGAAGGTATTAGACCTCTCTGGACCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAGACTGC	605
Qy			
	541	GAGAAGGTATTAGACCTCTCTGGACCCCTGCTTCGGGAGACCTGGTAAATCCGAGAAGACTGC	600
Db			
	606	CGGGGGAATATCTGATTTCCGGGTCTCTCCAGAAGCCCATCAGTAGCTTTTGCTGATTTTT	665
Qy			
	601	CGGGGGAATATCTGATTTCCGGGTCTCTCCAGAAGCCCATCAGTAGCTTTTGCTGATTTTT	660
Db			
	666	GAGCGGCATCTCTCGCCAGCCAGCTCGAAATCGGACTGTAGAGACCCCGGCTCAACAG	725
Qy			
	661	GAGCGGCATCTCTCGCCAGCCAGCTCGAAATCGGACTGTAGAGACCCCGGCTCAACAG	720
Db			
	726	CGCTCTCCCGCAGTCACTGTGTCTCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCC	785
Qy			
	721	CGCTCTCCCGCAGTCACTGTGTGTCTCTGCTCAAGGTGGACACAGCGGGAACGTTTGGCC	780
Db			
	786	CCAATTTCCGACGAGAGGGAAAAAATCTTACTGATTTGACTTTGGCTGGGTTCAGAGACCAAC	845
Qy			
	781	CCAATTTCCGACGAGAGGGAAAAAATCTTACTGATTTGACTTTGGCTGGGTTCAGAGACCAAC	840
Db			
	846	CGGGCGCACAGGCACCAAGGGCTTCGGCTAAAGAGAGTGAGAGCATCAACACCTCCCTG	905
Qy			
	841	CGGGCGCACAGGCACCAAGGGCTTCGGCTAAAGAGAGTGAGAGCATCAACACCTCCCTG	900
Db			
	906	TTTGTCTCTGGGC-AAAGTGGTAGATGGCTCAATCAGGGCTCTCCCTCTGTGTACCTTAT	962
Qy			
	901	TTTGTCTCTGGGC-AAAGTGGTAGATGGCTCAATCAGGGCTCTCCCTCTGTGTACCTTAT	958
Db			
	RESULT 12		
	BM910587	1031 bp mRNA linear EST 12-MAR-2002	
	LOCUS	AGENCOURT_6702754 NIH_MGC_98 Homo sapiens cdna clone IMAGE:5454101	
	DEFINITION	S., mRNA sequence.	
	ACCESSION	BM910587	
	VERSION	BM910587.1 GI:19360966	
	KEYWORDS	EST.	
	SOURCE	Homo sapiens (human)	
	ORGANISM	Homo sapiens	
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
	TITLE	1 (bases 1 to 1031)	
	JOURNAL	NIH-MGC http://mgi.nci.nih.gov/	
	COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs@rmail.nih.gov Tissue Procurement: ATCC cdna Library Preparation: Rubin Laboratory DNA Sequencing by: Agencourt Bioscience Corporation Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLCMI948 row: 1 column: 06 High quality sequence stop: 679. Location/Qualifiers 1..1031 /organism="Homo sapiens"	
	FEATURES		
	source		

Db 781 CTTATCGGACAGCAAGCTCACTCGCTATTGACAGACTCTCTGGTGGCTCAGCCAC 840  
 Qy 1017 AGTATCTTATTGCCAATTCATGCGCTCAGAGAGCTTCTACCTAGACACAGTCTCCGCA 1076  
 Db 841 AGTATCTTATTGCCAATTCATGCGCTCAGAGAGCTTCTACCTAGACACAGTCTCCGCA 900  
 Qy 1077 CTCACTTTGTCTGCCAGG-TCCAAGGAGGTGATCAATCGGCTTTT 1121  
 Db 901 CTCACTTTGTCTGCCAGGTTCCAGGAGGTGATCAATCGGCTTTT 946

RESULT 13  
 AL559769  
 LOCUS  
 DEFINITION AL559769 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA  
 clone CS0DG005YF11 5-PRIME, mRNA sequence.

ACCESSION AL559769  
 VERSION AL559769.3 GI:46185141  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 999)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:31283900.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 7766.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna/s=CS0DG005YF11&c=7766.r.  
 Location/Qualifiers

FEATURES  
 source  
 1. 999  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="caxon:9606"  
 /clone="CS0DG005YF11"  
 /tissue\_type="B CELLS (RAMOS CELL LINE)"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

ORIGIN  
 Query Match 59.9%; Score 921.4; DB 1; Length 999;  
 Best Local Similarity 96.8%; Pred. No. 7e-233;  
 Matches 913; Conservative 22; Mismatches 8; Indels 0; Gaps 0;

Qy 6 GCCGGCGCGCTCGACGACGAGCGCGAGATGGCGGAGCTTCAGCGCGCG 65  
 Db 41 GCCGGCGCGCTCGACGACGAGCGCGAGATGGCGGAGCTTCAGCGCGCG 100  
 Qy 66 ATCTCAGAGCTGGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGCTCCACCTCCA 125  
 Db 101 ATCTCAGAGCTGGTGGCTGTGGCTAAGCAAGATTGGAGCTACTCGTCGCTCCACCTCCA 160  
 Qy 126 GCTCGCTAAGGCTGGCTGTGGCTACTCGGCGCATTTGTGGATGGAAAGCGGAGCAAGT 185  
 Db 161 GCTCGCTAAGGCTGGCTGTGGCTACTCGGCGCATTTGTGGATGGAAAGCGGAGCAAGT 220

Qy 186 GATCCCCCTGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 245  
 Db 221 GATCCCCCTGTGTGCGGGCATGACAGCTGCTCTCTAGAGATTGCTAACTGGAGGAAC 280  
 Qy 246 CACGAGAGACTCTCAATACAGTTTGTATGCTCTTCTATGGGAGAGAGTACTCAGCAG 305  
 Db 281 CACGAGAGACTCTCAATACAGTTTGTATGCTCTTCTATGGGAGAGAGTACTCAGCAG 340  
 Qy 306 GACATCTATGACAGTTTGTATGCTCAGTCCAGCCATCTAAGGCACTTGTCTGAAGGGCAGAAATGCC 365  
 Db 341 GACATCTATGACAGTTTGTATGCTCAGTCCAGCCATCTAAGGCACTTGTCTGAAGGGCAGAAATGCC 400  
 Qy 366 AGTGTGCTTGTCTATGACCCACAGGAGCTGGGAGAGCGCACACAATGCTTGGGACGCCA 425  
 Db 401 AGTGTGCTTGTCTATGACCCACAGGAGCTGGGAGAGCGCACACAATGCTTGGGACGCCA 460  
 Qy 426 GAGCAACTCTGGGTGTATCCCGCGGCTCTATGACCTCTTCTGACCTCTGACGCTCACAAGGGAGAG 485  
 Db 461 GAGCAACTCTGGGTGTATCCCGCGGCTCTATGACCTCTTCTGACGCTCACAAGGGAGAG 520  
 Qy 486 GGTCCGAGGCGCGCCATGCGCTTCTGTCAACATGTCTTACCTAGAGATCTACCAAG 545  
 Db 521 GGTCCGAGGCGCGCCATGCGCTTCTGTCAACATGTCTTACCTAGAGATCTACCAAG 580  
 Qy 546 GAGAAGGTATTAGACCTCTCTGACCCCTGCTTCCGAGAGACCTGGTAAATCCGAGAACTGTC 605  
 Db 581 GAGAAGGTATTAGACCTCTCTGACCCCTGCTTCCGAGAGACCTGGTAAATCCGAGAACTGTC 640  
 Qy 606 CGGCGGAATATCTGATTTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 665  
 Db 641 CGGCGGAATATCTGATTTCCGGGTCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTTT 700  
 Qy 666 GAGCGGCACTTCTTCCGAGCCAGTCCGAATCGGAGCTGTAGGACCCCGCTCAACCCAG 725  
 Db 701 GAGCGGCACTTCTTCCGAGCCAGTCCGAATCGGAGCTGTAGGACCCCGCTCAACCCAG 760  
 Qy 726 CGCTCTCTCCGAGCTCATGTCTCTCTCAAGGTGGACCGGAGAGCTTTGGCC 785  
 Db 761 CGCTCTCTCCGAGCTCATGTCTCTCTCAAGGTGGACCGGAGAGCTTTGGCC 820  
 Qy 786 CCATTTCGCGAGGAGGAGGAAACTCTTACCTGATTTGCTGGTGTGGGTCTAGAGGACAAAC 845  
 Db 821 CCATTTCGCGAGGAGGAGGAAACTCTTACCTGATTTGCTGGTGTGGGTCTAGAGGACAAAC 880  
 Qy 846 CGGCGCAGGCAACAGGCGCTTCCGCTAAAGAGAGTGGAGGCATCAACACCTCCCTG 905  
 Db 881 CGGCGCAGGCAACAGGCGCTTCCGCTAAAGAGAGTGGAGGCATCAACACCTCCCTG 940  
 Qy 906 TTTGTCTCTGGGCAAGGTGTAGATGCGCTGAATCAGGGCCTCC 948  
 Db 941 TTTGTCTCTGGGCAAGGTGTAGATGCGCTGAATCAGGGCCTCC 983

RESULT 14  
 BX376199

LOCUS BX376199 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 DEFINITION CDNA clone CS0DC024YA10 5-PRIME, mRNA sequence.

ACCESSION BX376199  
 VERSION BX376199.2 GI:46616791  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1094)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)

JOURNAL On May 8, 2003 this sequence version replaced gi:30452476.  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 7766.r

For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?rs=CS0DC024BA05QPl&c=7766.r>.

## FEATURES

Location/Qualifiers  
 1. .1094  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DC024FA10"  
 /tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 59.9%; Score 921.4; DB 5; Length 1094;  
 Best Local Similarity 99.2%; Pred. No. 7.2e-233;  
 Matches 963; Conservative 4; Mismatches 0; Indels 4; Gaps 4;

QY 6 GCCGCGCGCGCTCGACGACGAGAGCGGACGAGATGCGCGAGCTTCAGCGG-CGGC 64  
 DB 12 GCCGCGCGCGCTCGACGACGAGAGCGGACGAGAGCGGCGGAGCTTCAGCGGCGCGC 71  
 QY 65 GATCTCAGGAGCTGCTCGCTGCTGCGCTAAGCAGAGATTGGAGTACTCGTCCACCTCC 124  
 DB 72 GATCTCAGGAGCTGCTCGCTGCTGCGCTAAGCAGAGATTGGAGTACTCGTCCACCTCC 131  
 QY 125 AGCTCGCGTAAGGCTGCTGCGACTCGCGCCATTGTGGATGGAACACGCGGAGCAAG 184  
 DB 132 AGCTCGCGTAAGGCTGCTGCGACTCGCGCCATTGTGGATGGAACACGCGGAGCAAG 191  
 QY 185 TGATCCCCCTGTGTCGGGGCATGAGACGCTGCTCTAGAGATTGCTAACTGGAGGAA 244  
 DB 192 TGATCCCCCTGTGTCGGGGCATGAGACGCTGCTCTAGAGATTGCTAACTGGAGGAA 251  
 QY 245 CCACGAGGACTCTCAATACAGATTGATGCTTCTATGCGGAGGAGTACTCAGCA 304  
 DB 252 CCACGAGGACTCTCAATACAGATTGATGCTTCTATGCGGAGGAGTACTCAGCA 311  
 QY 305 GGACATCTATGAGGTTTCACTGACGCCCATCTTAAGGCACTTCTGGAAGGCGAGATGC 364  
 DB 312 GGACATCTATGAGGTTTCACTGACGCCCATCTTAAGGCACTTCTGGAAGGCGAGATGC 371  
 QY 365 CAGTGTGCTTGCTATGACCCACAGAGCTGGGAGAGCCACACAAATGCTGGGCGAGCC 424  
 DB 372 CAGTGTGCTTGCTATGACCCACAGAGCTGGGAGAGCCACACAAATGCTGGGCGAGCC 431  
 QY 425 AGAGCAACCTGGGCTGATCCCGGGCTCTCATGGACCTCTCAGCTCACAAGGAGGA 484  
 DB 432 AGAGCAACCTGGGCTGATCCCGGGCTCTCATGGACCTCTCAGCTCACAAGGAGGA 491  
 QY 485 GGGTCCGAGGCGCGGCATATGGGCCCTTTCTGTCACCATGCTTACCTAGAGATCTACCA 544  
 DB 492 GGGTCCGAGGCGCGGCATATGGGCCCTTTCTGTCACCATGCTTACCTAGAGATCTACCA 551  
 QY 545 GGAGAGGATTAGACTCTGACCCCTGCTTGGAGAGCTGGTAAATCGGAGAACTG 604  
 DB 552 GGAGAGGATTAGACTCTGACCCCTGCTTGGAGAGCTGGTAAATCGGAGAACTG 611  
 QY 605 CCGGGGAATATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTT 664  
 DB 612 CCGGGGAATATCTGATTCGGGCTCTCCAGAGCCCATCAGTAGCTTTGCTGATTT 671  
 QY 665 TGAGCGGCACTTCTCCGACCCAGTCGAAATCGGACTGTAGGAGCCACCGGCTCAACCA 724

DB 672 TGAGCGGCACTTCTCCGACCGAGTCGAATCGGACTGTAGGAGCCACCCGGCTCAACCA 731  
 QY 725 GCCTCTCTCCGAGCATGCTGTGCTCTCTGTCAGGTGGACACGCGGGAACGTTTGGC 784  
 DB 732 GCCTCTCTCCGAGCATGCTGTGCTCTCTGTCAGGTGGACACGCGGGAACGTTTGGC 791  
 QY 785 CCCATTTCCGACGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTCAGAGGACAA 844  
 DB 792 CCCATTTCCGACGAGAGGGAACCTCTACCTGATTGACTTGGCTGGGTCAGAGGACAA 851  
 QY 845 CCGGCGCACAGGCAACAAGGGCTTCGGCT-AAAAGAGAGTGAGGACCATCAACACCTCCC 903  
 DB 852 CCGGCGCACAGGCAACAAGGGCTTCGGCTAAAAGAGAGTGAGGACCATCAACACCTCCC 911  
 QY 904 TGTGTTGCTCTGGGCAAGGTGATAGTCGCTGAATCAGGCGCTCCCTCTGCTGTTTATC 963  
 DB 912 TGTGTTGCTCTGGGCAAGGTGATAGTCGCTGAATCAGG-CCTCCTCGTGTA-CTTATC 969  
 QY 964 GGGACAGCAAG 974  
 DB 970 GGGACAGCAAS 980

## RESULT 15

BM546877

## LOCUS

BM546877 1019 bp mRNA linear EST 20-FEB-2002

## DEFINITION

AGENCOURT\_6491037 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5723623

## ACCESSION

BM546877

## VERSION

BM546877.1 GI:18780186

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 1019)

## AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

## TITLE

Unpublished (1999)

## JOURNAL

Contact: Robert Strausberg, Ph.D.

## COMMENT

Email: cgabbs@mail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12711 row: j column: 08

High quality sequence stop: 688.

## FEATURES

Location/Qualifiers

1. .1019

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5723623"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_125"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;

Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool

of three ovaries, from females ranging in age from 38 to

49 yo. Library is oligo-dT primed and directionally cloned

(EcoRV site is destroyed upon cloning). Average insert

size 2.1 kb, insert size range 1-3.5 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 036."

## ORIGIN

Query Match 59.1%; Score 909; DB 4; Length 1019;  
 Best Local Similarity 97.5%; Pred. No. 1.4e-229;  
 Matches 965; Conservative 0; Mismatches 21; Indels 4; Gaps 4;

Qy 21 ACGCAGCAGAGCGGCGACGCGAGATGCGCGCAGCTTCCAGCGCGCGCGATCTCAGGAGCTGGT 80  
 Db |||||  
 1 ACGCAGCAGAGCGGCGACGCGAGATGCGCGCAGCTTCCAGCGCGCGCGATCTCAGGAGCTGGT 60  
 Qy 81 CGCTGTGGCTAAGCAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAAGGTTG 140  
 Db |||||  
 61 CGCTGTGGCTAAGCAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAAGGTTG 120  
 Qy 141 GCTGTGGCTAAGCAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAAGGTTG 200  
 Db |||||  
 121 GCTGTGGCTAAGCAGATTGGAGCTACTCGTGGTCCACCTCCAGCTCGCGTAAAGGTTG 180  
 Qy 201 CGGCGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCTC 260  
 Db |||||  
 181 CGGCGCATGGACAGCTGCTCTAGAGATTGCTAACTGGAGGAACCAACAGGAGACTCTC 240  
 Qy 261 AAATACCAAGTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGGT 320  
 Db |||||  
 241 AAATACCAAGTTGATGCCCTTCTATGGGAGAGGAGTACTCAGCAGGACATCTATGCAGGT 300  
 Qy 321 TCAGTGCAGCCATCCTAAGGACCTTGTGGAAGGCGAGATGCCAGTGTGCTTGCCTAT 380  
 Db |||||  
 301 TCAGTGCAGCCATCCTAAGGACCTTGTGGAAGGCGAGATGCCAGTGTGCTTGCCTAT 360  
 Qy 381 GGACCCACAGGAGCTGGGAGACGACACAATGCTGGGCGAGCCAGAGCAACCTGGGGTG 440  
 Db |||||  
 361 GGACCCACAGGAGCTGGGAGACGACACAATGCTGGGCGAGCCAGAGCAACCTGGGGTG 420  
 Qy 441 ATCCGGGGGCTCTATGAGCTCTGTCAGCTCACAGGGAGAGGGTCCGAGGGCGG 500  
 Db |||||  
 421 ATCCGGGGGCTCTATGAGCTCTGTCAGCTCACAGGGAGAGGGTCCGAGGGCGG 480  
 Qy 501 CCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGATATTAGAC 560  
 Db |||||  
 481 CCATGGGCGCTTCTGTCACCATGCTTACCTAGAGATCTACAGGAGAGGATATTAGAC 540  
 Qy 561 CTCTGGACCTGCTCGGAGACCTGGTAAATCCGAGAGACTGCGGGGGAATATCCTG 620  
 Db |||||  
 541 CTCTGGACCTGCTCGGAGACCTGGTAAATCCGAGAGACTGCGGGGGAATATCCTG 600  
 Qy 621 ATTGGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGTAGCGGCACTTCTG 680  
 Db |||||  
 601 ATTGGGGTCTCTCCAGAGCCCATCAGTAGCTTGTGATTTTGTAGCGGCACTTCTG 660  
 Qy 681 CCAGCCAGTGAATCGGAGCTGAGGCGCACCGGCTCAACAGCGCTCCTCCCGCAGT 740  
 Db |||||  
 661 CCAGCCAGTGAATCGGAGCTGAGGCGCACCGGCTCAACAGCGCTCCTCCCGCAGT 720  
 Qy 741 CATGCTGTCTCTGTCAGGTCAGGTCAGCGGGAACGTTTGG-CCCCATTTCCGCCAGCG 799  
 Db |||||  
 721 CATGCTGTCTCTGTCAGGTCAGGTCAGCGGGAACGTTTGGCGCCCATTTCCGCCAGCG 780  
 Qy 800 AGAGGGAAACTCTACCTGATTTGATCTTGGCTGGGTCTAGAGCAACCGGCGCAAGGCAA 859  
 Db |||||  
 781 AGAGGGAAACTCTACCTGATTTGATCTTGGCTGGGTCTAGAGGCAACCGGCGCAAGGCAA 840  
 Qy 860 CAAGGGCTTCGGCTAAAGAGCTGGAGCCATCAACACCTCCTCTGTTGCTTGGGCAA 919  
 Db |||||  
 841 CAAGGGCTTCGGCTAAAGAGCTGGAGCCATCAACACCTCCTCTGTTGCTTGGGCAA 900  
 Qy 920 AGTGGTAGCTGCTGAATCAGGGCCTCCCTCTGTTGTA-CTTATCGGACAGCAAG-CTC 977  
 Db |||||  
 901 AGTGGTAGCTGCTGAATCAGGGCCTCCCTCTGTTGTAACCTTAATCGGACAGCAAGCTC 960  
 Qy 978 ACTGCCTATTGCA-GGACTCTGGGTGG 1006  
 Db |||||  
 961 ACTGCCTATTGCAAGGACTCTCTGGGTGG 990

This Page Blank (uspto)